A UNITED STATES DEPARTMENT OF COMMERCE PUBLICATION

Regional Economic Development Research Conference

April 19, 1972

PROCEEDINGS

U.S. DEPARTMENT OF COMMERCE

Economic Development Administration

Regional Economic Development Research Conference

PROCEEDINGS

April 19, 1972

U.S. DEPARTMENT OF COMMERCE
Peter G. Peterson, Secretary

Robert A. Podesta, Assistant Secretary
for Economic Development

The Regional Economic Development Research Conference was held at the National Bureau of Standards, Administration Building, Green Auditorium, Gaithersburg, Maryland.

EDA OER-72-004

CONTENTS

Page

Agenda

Opening Remarks
Robert A. Podesta 1

Conference Orientation and Overview
Samuel M. Rosenblatt 3

Regional Growth Centers: A Status Report
Brian J. L. Berry 9

Behavioral Models of Inter-Regional Migration
Irving N. Fisher 47

National Growth Policy and the Environmental Effects of Cities
George S. Tolley..... 55

Estimating the Regional Impact of Federal Income Maintenance Programs
T. Paul Schultz..... 85

Assessing the Differential Regional Consequences of Federal Tax-Transfer Policy
Stephen P. Dresch 91

Policy Implications of Intermetropolitan Migration Flows
William Alonso 143

Summary - Concluding Remarks
William W. Blunt, Jr. 153

Agenda

8:30 a.m.	Registration and Coffee	
9:00 a.m.	Welcome . . . Opening Remarks	Robert A. Podesta Assistant Secretary for Economic Development U.S. Department of Commerce
9:15 a.m.	Conference Orientation and Overview	Samuel M. Rosenblatt Director, Office of Economic Research, Economic Development Administration
9:30 a.m. to 11:00 a.m.	Population Mobility and Urban Concentrations	
	Chairman	Dr. Marina Whitman Council of Economic Advisers
	Regional Growth Centers—A Status Report	Prof. Brian Berry University of Chicago
	Behavioral Models of Interregional Migration	Irving Fisher The New York City Rand Institute
	Economic Effects of Population Distribution Policies	Prof. George Tolley University of Chicago
11:00 a.m. to 11:20 a.m.	Coffee Break	
11:20 a.m. to 12:45 p.m.	Three Concurrent Workshops	
	Workshop 1—Lecture Room A Moderator	Prof. William Miernyk West Virginia University
	Workshop 2—Lecture Room B Moderator	Prof. Benjamin Chinitz Brown University
	Workshop 3—Lecture Room C Moderator	Prof. Leon Moses Northwestern University

1:00 p.m. to 2:00 p.m.	LUNCH	
2:00 p.m. to 3:30 p.m.	Federal Policies and Regional Development	
	Chairman	Raymond J. Waldmann Assistant to the President
	Regional Impact of Income Transfer Programs—The Family Assistance Plan	T. Paul Schultz The Rand Corporation
	Regional Effects of Federal Transfer Payments—Revenue Sharing and Grants-in-Aid	Stephen Dresch National Bureau of Economic Research
	Policy Implications of Inter-Metropolitan Population Flows	Prof. William Alonso University of California, Berkeley
3:30 p.m. to 3:50 p.m.	Coffee Break	
3:50 p.m. to 5:15 p.m.	Three Concurrent Workshops	
	Workshop 4—Lecture Room A Moderator	Michael Koleda National Planning Association
	Workshop 5—Lecture Room B Moderator	Joel Bergsman The Urban Institute
	Workshop 6—Lecture Room C Moderator	Harold M. Hochman The Urban Institute
5:15 p.m. to 6:00 p.m.	Summary of Workshop Sessions	
	Chairman	William W. Blunt, Jr. Deputy Assistant Secretary for Economic Development Economic Development Administration

CONFERENCE PROCEEDINGS

Opening Remarks of Robert A. Podesta

We are glad to welcome you to this conference, and to have the opportunity to exchange views with you on the results of some of the research that has been going on under EDA sponsorship during the past year. Six of our distinguished research scholars will present papers on work they have in progress. The workshops following their presentations will provide a chance for informal questions and discussions in smaller groups. We hope in this way to benefit from your reactions and responses to the ideas they develop.

We are honored to have Dr. Marina Whitman, a member of the Council of Economic Advisers, as Chairman of the morning session - Population Mobility and Urban Concentrations, and Mr. Raymond J. Waldman, Assistant to President Nixon, as Chairman of the afternoon session - Federal Policies and Regional Development. We believe that the presence of these Chairmen at our conference, as well as all our other distinguished guests, is indicative of the great interest and concern throughout the government in the questions to be discussed today.

The wide variety of subjects to be covered at the conference reflects EDA's belief that the agency's research program can provide a firm foundation for the formulation of a national growth policy. We look to this conference as one block in the construction of this foundation. This conference is also one element in our program to disseminate more widely the results and findings of our research efforts. It is through such activities as this conference that these findings can be digested and analyzed by others and put to the test of policy and program application. We are hopeful that all of you will be stimulated to make full use of those aspects of our research program that are applicable to your respective areas of responsibility. We also want all of you to let us know how we can make this research effort even more useful. We look forward to a day of productive discussion.

CONFERENCE ORIENTATION AND OVERVIEW

SAMUEL M. ROSENBLATT

In the U.S. in the midst of our generally prosperous national economy there have almost always been regions and areas in the country that have lagged behind the rest of the nation economically. For example, to take one illustration from the earlier portions of our history, there is the case of Gloversville, New York--an area that initially grew because of the nearness of deer skins for use in glove manufacturing--but which then had to adjust as this raw material became exhausted. Or to skip ahead to the turn of the 19th century and the beginning of this century, the nation always experienced hard conditions in what today we refer to as "ethnic centers" of our larger, mainly eastern seaboard centers, but which were then simply called ghettos. In the 1920's New England experienced the effects of a textile depression and many of our agricultural areas suffered from low and declining agricultural prices. Of course in the 1930's the entire nation was affected though certain areas and regions were much harder hit than others--for example, the dust bowl.

However, it is only in recent years that we have begun to pay any concerted attention to this type of regional problem rather than simply accepting it as a fact of life and one of the side effects of overall national growth. Today, we are painfully aware of the problems of eastern Kentucky, of the so-called Black Belt that runs through Georgia, Alabama and Mississippi, of the Mexican-American border area, of the urban ghettos of many of our large cities, and of the poverty that exists on most Indian reservations. As a nation we are beginning to come to grips with these problems and develop solutions for them. However, this effort is a fairly recent phenomenon and an objective observer might well say that we have really only begun the struggle. I would tend to agree with this conclusion.

First, it is only recently that the American people have felt reasonably comfortable about the overall prospects for the national economy. We have gotten over our fears of future deep seated depressions, with its attendant high levels of unemployment and intense and widespread personal distress, and have come to rely on the ability of our economy, with the assistance of the fiscal and monetary tools of the Federal government, to ride out and adjust to economic change. That doesn't mean that life is a bed of roses for the Council of Economic Advisers these days and that all the problems and uncertainties associated with inflation and balance of payments difficulties will disappear with a wave of one's hand, but it does seem to mean that, as a people, we now anticipate that changes in our economy in the future will take place within narrower bounds. Given this situation, as a nation, we can direct more attention and resources towards the regional distribution of economic opportunities and economic change.

A second factor that affects this concern, and it sounds almost platitudinous to say this again -- it has been repeated time and again -- is that we clearly recognize the close interrelations and connections between the problems of the different parts of our country as never before. The rural outmigration of the 40's and 50's and the metropolitanization of our population in the 60's are facts of life and the interconnections between them are very clear.

A third factor, which really highlights the second, was the outbreak of violence in the urban, mostly black, ghettos, our major metropolitan areas. These outbreaks certainly sensitized the nation to the kinds of problems that existed in these places, as well as elsewhere in the nation.

A final factor that I think should be included was the experience the U.S. gained as a result of its assistance programs to Western Europe and the less developed nations of the world after World War II. These made us aware of the problems of underdevelopment and also provided us with first hand knowledge of how to run programs of economic development.

Today we have a number of specifically regional development programs, such as EDA, the Appalachian Regional Commission, and the Title V Regional Commissions. We also have a whole galaxy of other programs, which though not specifically regional, clearly have regional implications. These include manpower and vocational training programs, housing programs, highway construction, poverty programs, etc. These efforts are a start but I think that as we get more deeply involved, we will begin to uncover more and more of the complexities and difficulties associated with trying to solve these regional and urban problems.

What are some of these complexities? Well for one thing a regional development problem is not single faceted. It shows itself in many ways. For ease of presentation these problems may be categorized as individual problems, areal problems, and industrial problems. The solutions for any one set of problems may not add to, and indeed may detract from, a solution for another set of problems. Indeed the administration of regional development programs and policies, by their very nature, involves compromises among a set of potentially competing objectives.

As regards these categories of problems, for the individual, limited local employment opportunities may mean either remaining in an area and subsisting at the modest standard of living available locally, or else choosing to migrate to some other more prosperous region to improve this standard. For depressed or underdeveloped areas the development problem may be tied to an inadequate or declining resource base--natural, man-made, or human--whereas for growing and already prosperous areas the problem may be their ability to absorb new resources without distorting or undermining this growth. The problem for industries is associated with the dynamic nature of demand, resource depletion, and technological change. Old industries die and new ones spring up to compete with existing industries. Regional concentrations of affected industries result in areal problems and these have obvious implications for individuals.

EDA's Office of Economic Research has structured its research program in an attempt to broaden our understanding of these problems, highlight their interrelatedness, as well as put forth some solutions. In so doing, we hope to provide a basis for developing more effective regional development programs and policies in the future. Without trying to steal the thunder of the speakers of the day, let me try to outline and highlight some of the areas of research that we have pursued.

Regarding the individual who is part of an environment of economic decline or stagnation, his problems really demonstrate the fundamental issue and dilemma of regional development, namely, whether to bring employment opportunities to his locale or to move him to where the new employment opportunities are likely to exist. Without trying to resolve this hoary chestnut at this conference, let me simply note that present U.S. policy is to have a little of both.

Within this context we have supported a considerable amount of research effort on detailing interregional migration trends

- To try to identify where people who migrate come from and where they go--
- Whether they make this move in one full swoop or as a series of shorter moves--
- What happens to migrants once they get to their destinations--how do they fare economically as compared with their newly acquired neighbors as well as those they left behind?
- What kinds of places are migrants most likely to go to and what places are they most likely to leave?
- What are the characteristics of migrants--their education, sex, race, etc.?

Our research grantees have made use of much of the available data on this question--the special tabulations of the population censuses--OEO's Survey of Economic Opportunity--and the 1 percent Continuous Work History Social Security sample. As regards this latter series, OER has also assisted in creating an operational file that has been used both by external research recipients and in house personnel.

For the individual who chooses not to migrate, for whatever reasons, there is the related set of problems of enhancing local or nearby employment opportunities. This involves both questions of improving the employability of the individual and questions of what it would take in the way of economic and social investments to increase the number and raise the quality of available economic opportunities in these places. OER has not expended a great deal of its resources on the first of these questions--we have left

this task pretty much to the Manpower Administration and OEO among others However, we have engaged in a considerable amount of effort on the second set. We have examined such questions as the appropriate role of infrastructure investment--both social and economic--to try to determine which types of overhead capital would be most productive in the way of creating job opportunities. We have supported work on the efficiency and responsiveness of local and regional capital markets and institutions, both public and private. In addition work has been produced describing some desirable forms, structures, and administrative techniques for organizing local economic development programs.

One final effort involves the effects on the individual of a declining community population base as regards the quality, availability, and cost of public services. As Professor Alonso will note in his remarks this afternoon the phenomenon of actual population decline is likely to affect many of our metropolitan areas in the 1970's to say nothing of the increasing number of smaller rural counties, many of which have already experienced this condition.

As is readily apparent the focus on the individual and the problems associated with providing him with economic opportunities blends in quite naturally with the spatial or geographic aspects of economic development. Regional and urban economics is unique in the general discipline of economics only in that it adds a spatial dimension to the general list of variables with which the science is concerned. Regional economists study not only the what, how, how much, and to whom questions regarding national production and income but add also the question of where. In this sense, physical place and space take on a critical importance. They also obviously take on added meaning as regards regional development programs and policies since the latter are concerned about where it is that new economic opportunities ought - might - can be created.

Within this context EDA's Office of Economic Research has devoted a great deal of effort and resources to try to clarify the fundamental question of the location of economic activity. Particular attention has been paid to the role that urban places play in regional development. We have tried to identify the impacts that individual central places have on their outlying hinterland areas. This is an especially important question to EDA since the agency has been legislatively mandated with the responsibility of conducting a portion of its regional development programs through the use of local growth centers. Growth centers are places that are growing or have, as a minimum, indicated some potential for future economic growth. Professor Berry's work, which has been seminal in this area, provides the basis for much of the analytical effort now being pursued on this question in and out of government.

The agency has also studied the distribution of intra-metropolitan economic activity, though not to the same extent as the inter-metropolitan and urban-regional questions. With the increase in agency involvement

with inner city and suburban type problems this question is likely to receive increased attention. We have sponsored some work on the role of commuting, however, to determine reasonable time and distance limitations on this process and to explore the use of commuting as a substitute for migration.

One final point in this general area concerns the matter of knowing just what is happening, in an economic sense, at the sub-national level. As anyone who has worked in regional and urban economics can tell you, one of the major stumbling blocks to better economic analysis and policy prescription has been the absence of good, reliable, timely and uniform data for small areas on which to base this analysis.

At the very outset of its history, the agency undertook to do something about this. It worked closely with the Bureau of Economic Analysis of the Department of Commerce to support the development of a series of county total and per capita income estimates comparable to the series that BEA had developed for SMSA's and states. BEA also developed, with EDA support, a companion series of employment estimates at the county level to go with its income series. Both series are now available on an annual basis approximately 18 months after the close of the year.

Turning now to our third and final category of regional development problems--the industrial category--OER has undertaken both basic and applied research on the question of industrial location. In the former area we have helped pull together an annotated bibliography of the literature on industrial location theory and practice, while in the latter area we are in the process of developing an industrial location planning system. This system, basically simple in concept, but highly involved and complex in development, will match the specific locational and resource demands of industry, identified down to the individual product class, with the resource availabilities of EDA growth centers and redevelopment areas. These latter are places of unique concern to the EDA. They are areas throughout the nation whose economies are lagging because of structural problems or because they are basically underdeveloped. This project is designed to assist these areas do a more effective job of identifying potential new industries, whose needs are most suited to the resources available in these places. The actual task of locating a specific firm or industry in the area is still left up to the local area residents and the private sector.

Another aspect of the industrial problem as it affects regional development that we have looked into concerns the relationships between economic development and the environment. In particular we are examining how a region of the country can anticipate, measure, and deal with the pollutant creating activities associated with a given set of projections regarding the future composition of its industrial activity.

I have not attempted to provide you with an exhaustive catalogue of

EDA sponsored research this morning. Rather, I have tried to give you something of the flavor of our research concerns. You've really only been exposed to the tip of the iceberg. In particular I have avoided any mention of the specifics of some of our policy related research since all of us will be dealing with some of these questions as the day progresses.

The principal purpose of this conference and its attendant workshops is to communicate to all of you concerned with regional and urban development the findings, as tentative as these might be, of some of the research sponsored by EDA. In turn we are hoping that the ideas that each of you brings to these discussions will better inform us so that in the future we may improve the usefulness and relevancy of the research that we undertake.

REGIONAL GROWTH CENTERS: A STATUS REPORT

BRIAN J. L. BERRY

What I have to contribute to this conference is very much a status report of research in mid-stream, suffering from delays in its promised completion by the lagging fourth-count delivery schedules of the 1970 census. Already, of course, many of the broad features of regional growth in the sixties have become clear and I will begin by talking about some of these. Following this prologue, I will devote the balance of the paper to a description of several facets of our analysis of regional growth centers, by means of an example. Although this analysis remains incomplete, I will try to provide some appreciation of the consistent nationwide picture of the U.S. urban hierarchy and gradients of urban influence that we will be delivering later this year, and the relationships of this picture to the growth center strategies that, quite rightly I feel, remain at the core of regional development policy.

The point of departure for our research is the concept of the country as a system of metropolitan labor markets - of employment centers and their tributary hinterlands - with what has generally been termed "the regional poverty problem" increasing to the periphery and reaching its maximum in the isolated regions beyond daily contact with the metropolitan system. Accordingly, the task of regional policy is seen as one of strengthening key urban centers that can bring growth to the peripheries and the pockets of isolation beyond by (1) strengthening the links between these centers and the metropolitan system, for example through transportation and communications programs, (2) strengthening the ties between the centers and their outlying local hinterlands, and (3) development of service delivery systems for both the centers and their regions, uniting them in a common and equal fabric of services.

In this framework, the problem for regional growth center research is seen as one of tracing the links between the metropolitan system, outlying centers, tributary hinterlands and poverty pockets; identifying within the existing urban system potential growth centers capable of exerting uplifting spread effects on their hinterlands; and of determining the links between metropolitan progress and hinterland development.

Any pursuit of such questions quickly reveals that no uniform set of national standards can apply; the regional expressions of growth and change are highly particularistic and full of local detail. Yet equally, a comparative base is required for identifying regional differences, and this is provided in the form of a nation-wide specification of the nature and extent

of metropolitan labor markets and tributary hinterlands, and a consistent national definition of the urban hierarchy, as well as in consistent means of applying gradient analysis to the determination of those centers capable of exerting spread effects of different degree beyond their county limits. Each of these is an ingredient of our research.

I am going to assume that the audience is familiar with my earlier work for the Economic Development Administration and related agencies, and in particular with the studies we completed following the 1960 census relating commuting patterns, zones of metropolitan influence, hierarchical diffusion, spread effects and regional growth in the decade of the 1950's. Let it also be clear right at the start that I make no apology for some of the apparently simple graphic devices I use in my research to illustrate gradients of influence, labor markets, hinterlands and spread effects. I have found by hard experience that the multiequation model, the big computer and the packaged program all too often tend to obfuscate when they should enlighten, veiling the gap between concept - however elegant - and reality, subordinating details of local variability to the simple-minded elegance of the imposed hypothesis and the regression line. Gradient analysis is performed more rapidly and efficiently by graphic means. I try to be continually mindful of Sir William Petty's admonition three centuries ago: "a judicious parallel may be drawn between the Body Natural and the Body Politic, and between the arts of preserving both in health and strength, and as it is anatomy is the best foundation of one, so also of the other: to practice upon the politic, without knowing the symmetry, fabric, and proportion of it, is as casual as the practice of old women." Some of the symmetry, fabric and proportion of regional growth centers in the 1960's are what our research is about, and what our graphics seek to portray.

Regional Growth: Nationwide Changes in the Sixties

That the regional contexts within which growth center strategies will have to work very considerably is unquestioned. This being so, I would like to begin by assuming that an audience at a Regional Economic Development Research Conference is quite familiar with the variety of such contexts, and especially with the population distribution and redistribution dynamics discussed in the second chapter of the Rockefeller Commission's report on Population and the American Future as well as in the Domestic Council's Report on National Growth, 1972. A few highlights should serve as reminders.

Both reports conclude that "population growth is metropolitan growth in the contemporary United States," but they also show that "population growth has been a dual process of concentration on a national scale and dispersion and expansion at a local level."

First, what were some of the facts of concentration? The total metropolitan population grew by 26 millions in the 1960s, one third due to territorial expansion of existing centers and two thirds from growth within

constant boundaries. Three quarters of the latter was natural increase, and 26 percent involved net migration from nonmetropolitan areas and from overseas. Meanwhile, nonmetropolitan populations declined, while the remaining people in such areas continued to concentrate in smaller urban centers, leaving less than ten millions on the farm.

What about dispersion and expansion at a local level? New roads and modern communications have enabled metropolitan areas to disperse in broad Daily Urban Systems. Figure 1 shows the centers of metropolitan status in 1960, according to our hierarchical scheme, and Figure 2 their commuting areas in the same year. The Daily Urban System idea is related to these commuting patterns, for what Figure 2 shows are, in effect, the geographical labor and housing markets of the metropoles - the areas linked by daily ebbs and flows of people, goods and ideas. More extensive than the SMSA's, these are the real geographical expressions of everyday interdependence, the "true" metropoles.

Within these systems, internal variability is increasing along racial, ethnic, social, economic, age and environmental lines. Accelerating dispersion meant in the 1960s that more massive depopulation occurred in the traditional central cities, increasingly abandoned by the more affluent majority to growing minority groups, than in declining rural counties. Figure 3 gives some sense of the mutually antagonistic residential choices of blacks and whites: the greater the growth of the ghetto, the greater the decline of the central city population. Territorially expanding suburbs and exurbia captured the metropolitan growth of the 1960s of all except the minorities. The affluent, the whites, new jobs, services, facilities and residences grew in increasingly more highly differentiated outlying settings, multi-nodal rather than core oriented, dispersed rather than centralized.

The internal problems resulting are profound - the geographical emergence of the Kerner Commission's "two societies" along central city lines; and the divergence between the "real city" - the functionally integrated Daily Urban System - and the legal entities that are supposed to govern it.

All of these things should be well known to this audience, and what they do is but provide context for our studies of growth centers in their regional context. The picture of urbanization in the 1950's was still one of the role of size, density, and heterogeneity as suggested by the Department of Agriculture's 1960 picture of urban orientation of counties shown in Figure 4. During the 1960s, the core cities of great size declined, however. The map of the areas growing more rapidly than the nation 1960-1970 (Figure 5) points clearly enough to the transformation taking place. Core-oriented metropoles merged into megalopoles and within megalopolitan regions growth was most rapid in the areas of lowest densities. Dispersion and extension characterized the major regions of national population concentration. The effective circulation territory of the Twin Cities increased from 1,000 to over 15,000 square miles in the last two decades, for example, with the greatest change coming in the second half of the period.

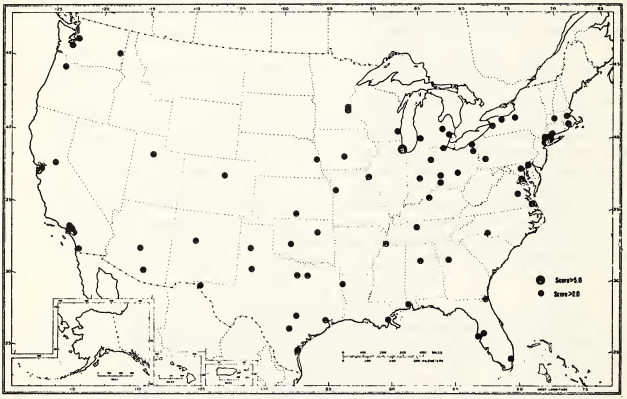


FIGURE 1: The Metropolitan Centers in 1960

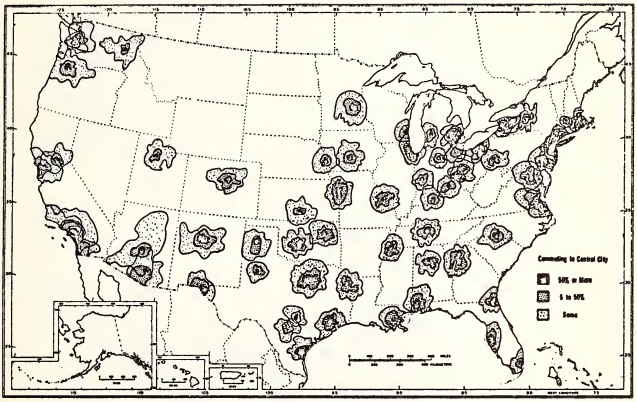


FIGURE 2: The Metropolitan Areas of 1960 - A Commuting Field Definition

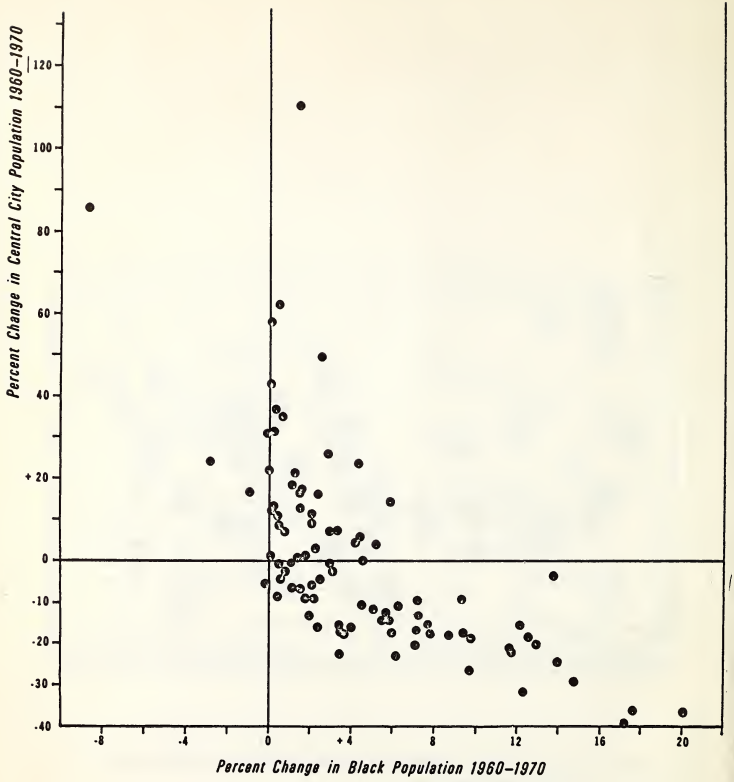


FIGURE 3: The Inverse Relationship Between Population Change in Central Cities and Increases in their Black Population Percentage, 1960-1970

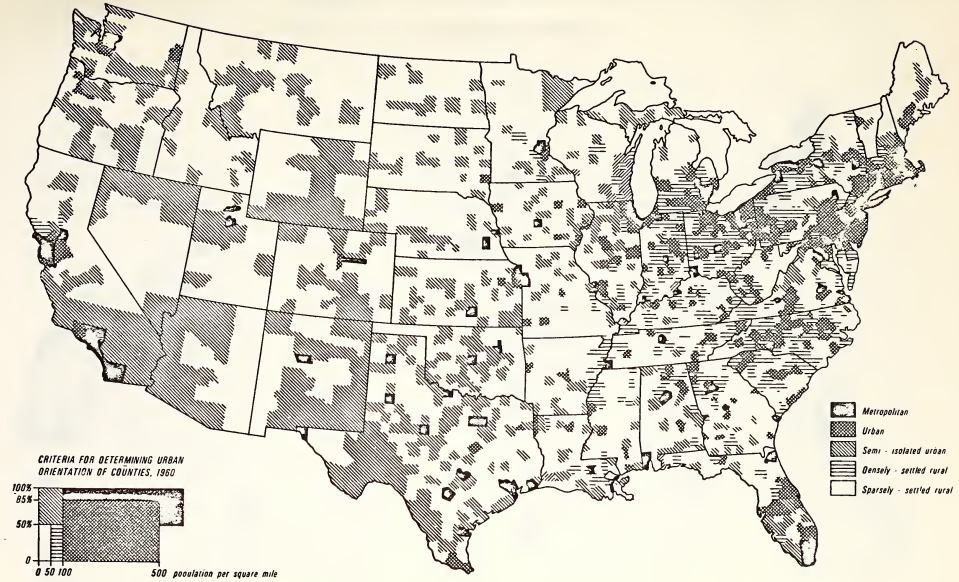


FIGURE 4: Urban Orientation of Counties in 1960, as Defined by the U.S. Department of Agriculture Using Criteria of Size and Density

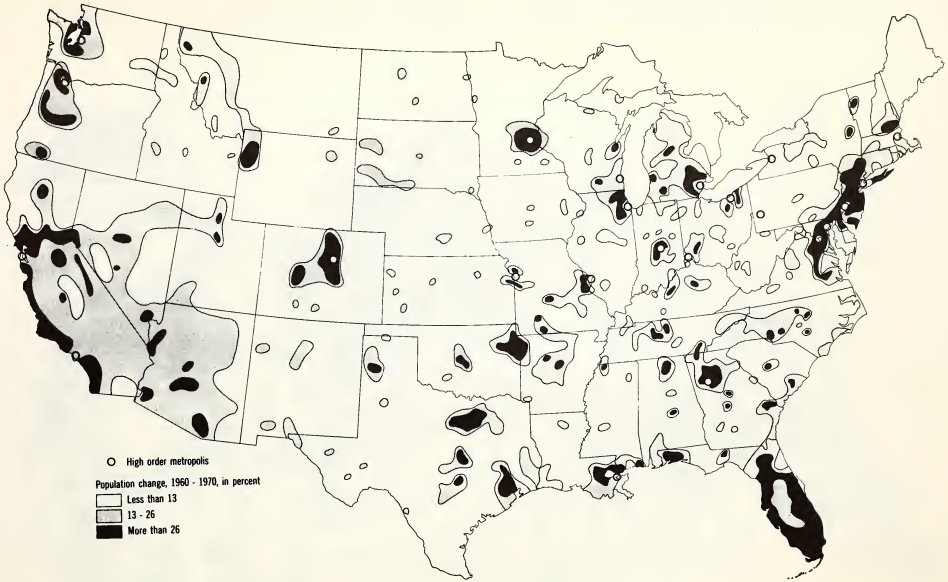


FIGURE 5: PATTERNS OF POPULATION CHANGE 1900-1970

J. R. BORCHERT'S 30 "HIGH ORDER" METROPOLES ARE IDENTIFIED

Yet all of these dynamics contain within them other features of interest to those concerned with regional development. Let us begin by agreeing that the OBE Economic Areas (Figure 6), based as they are upon the nation's labor markets in the 1960's (Figures 7 and 8), are indeed regional economic entities with a high degree of closure in the short-run. Then graph their population growth rate 1960-1970 against 1960 population size, as in Figure 9. A regression line would be positive. But more information is provided by dividing the array into size-classes and computing the median growth rate for each size-class, as well as the quartiles. The following conclusions then emerge:

1. The median population growth rate of successive size classes increased progressively with size to a population of 1,000,000, and stabilized thereafter at about the national growth rate. This is consistent with Thompson's ideas of the relationship of large size to stable self-generative growth.

2. The inter-quartile range was stable for size classes of less than 1,000,000, and above that point also for the lower quartile. However, the upper quartile was markedly greater than elsewhere in the size class 1,000,000-2,224,000, indicating an accelerated "take-off" of many economic areas of this size range in particular.

3. The median growth rate was negative in the smallest size class, as was the lower quartile in the size range 225,000-500,000, indicating that declining DUS's are disproportionately the smaller ones.

What accounts for deviations from these size-related growth trends? Some cross-classification helps here, as in Table 1, which compares size with economic base of the economic areas, as evidenced by sources of earnings in 1967, and with status of the areas in the nation's urban hierarchy, about which more will be said later.

Scan the columns of the table. Either the federal government or residentiary (nonbasic) activities, or a combination of the two, were the principal sources of earnings in the rapid-growth economic areas of the 1960-1970 decade. Together, the concentration of federal expenditures in particular places, and the rise to ascendency of the service sector, provided the propulsive sources of regional growth at rates greater than the nation, especially in intermediate-sized urban regions, in the decade. While the largest urban systems grew at about the national growth rate - itself fluctuating as a result of national fortunes and policies - the wider variations in growth performance of the smaller DUS's were the result, on the negative side, of a combination of increasing agricultural efficiency or resource depletion, and on the positive side of service-sector growth and the federal dollar.

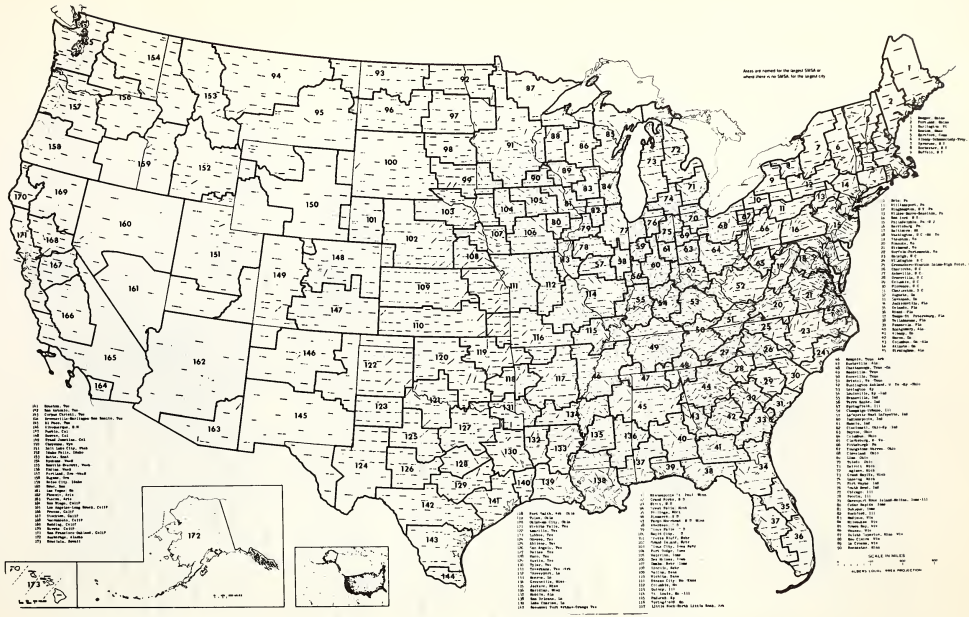


FIGURE 6 : THE OBE ECONOMIC AREAS

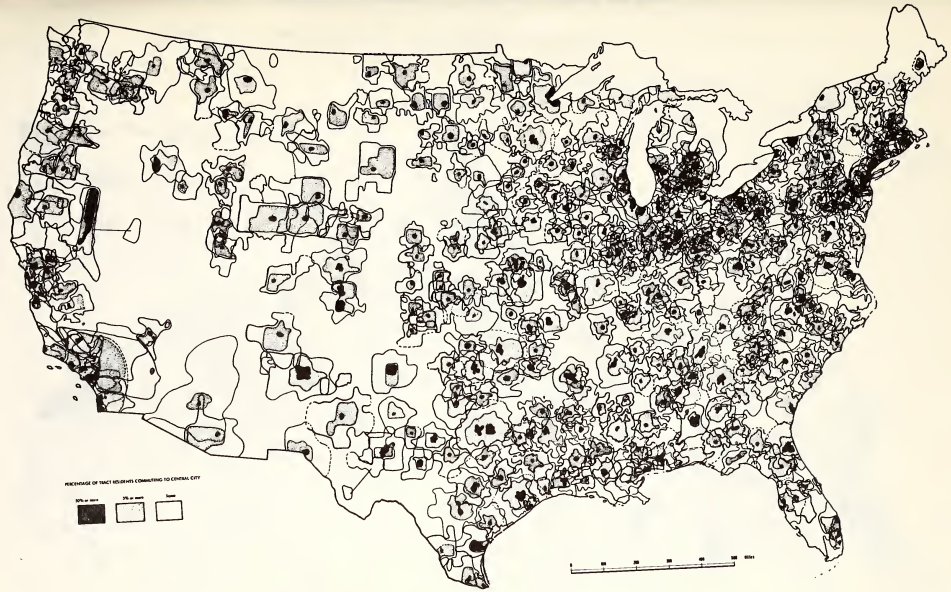


FIGURE 7 : THE NATIONAL PATTERN OF URBAN COMMUTING FIELDS IN 1960.

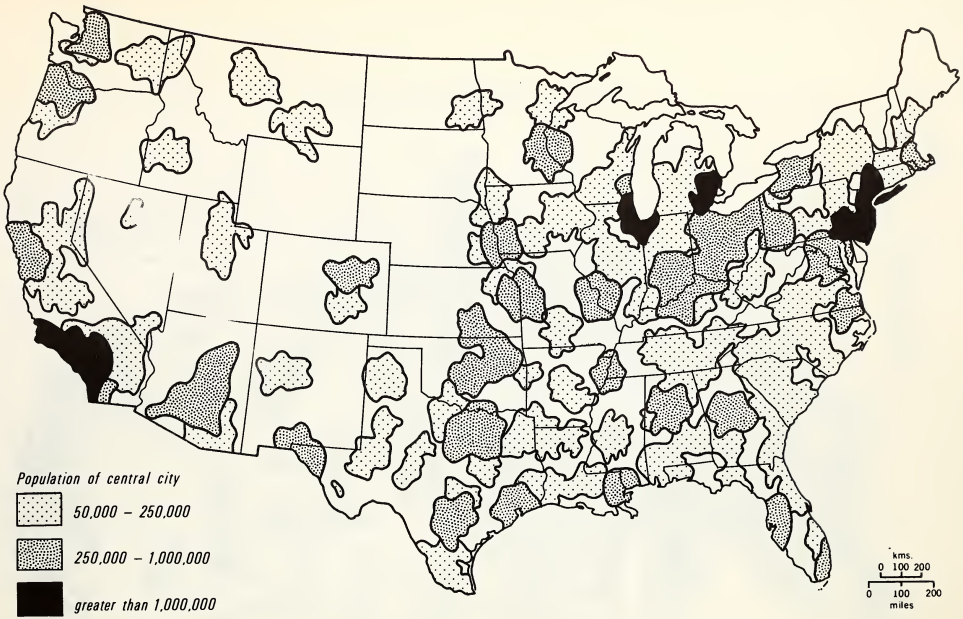


FIGURE 8 : AREAS WITHIN COMMUTING RANGE OF CENTRAL CITIES OF VARIOUS SIZES IN 1960

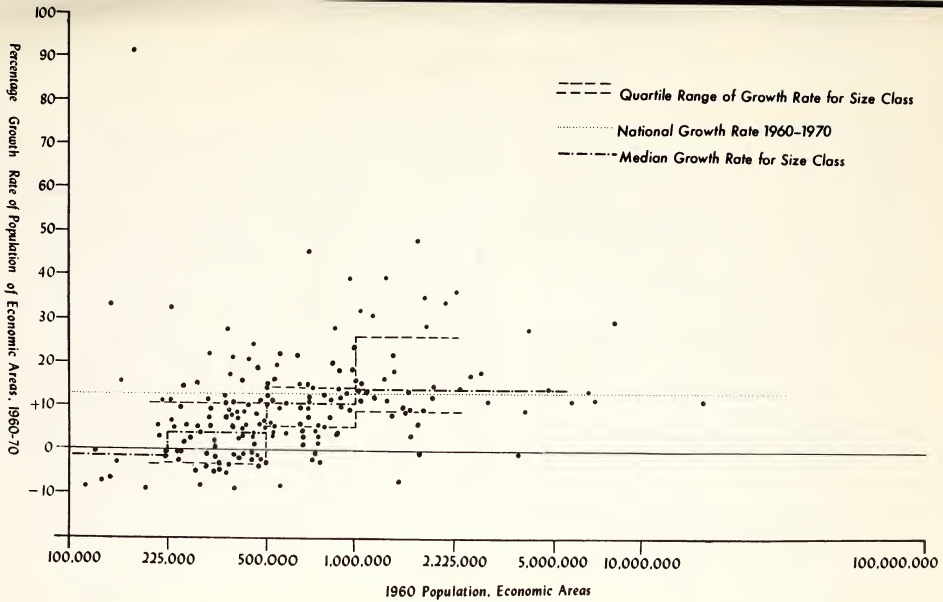


FIGURE 9 : GROWTH RATE OF THE ECONOMIC AREAS 1960-1970 ,

RELATED TO THEIR SIZE IN 1960

TABLE 1

MEDIAN GROWTH RATES OF VARIOUS TYPES OF ECONOMIC AREAS, 1960-70

POPULATION SIZE-CLASS OF ECONOMIC AREA IN 1970 (in thousands)	ECONOMIC AREA FOCUSSES UPON: CENTER OF METROPOLITAN STATUS and source of earnings of area was derived in 1967 from: >12.5% Federal Government	>60% Residentiary	Federal and Residentiary	Diversified Sources	CENTER OF LESS THAN METROPOLITAN STATUS and source of earnings of area was derived in 1967 from: >12.5% Federal Government	>60% Residentiary	Federal and Residentiary	Diversified	>25% Agriculture	Federal and Agriculture	>12.5 Mini
1. 100-225	a	a	a	a	15.3	-2.05	b	8.6	-8.05	b	b
2. 225-500	b	1.7	23.7	-3.0	9.7	4.9	b	4.0	-4.0	-2.5	-4.6
3. 500-1,000	7.2	11.0	18.2	15.5	10.7	12.0	b	9.6	b	b	b
4. 1,000-2,250	27.2	18.9	15.4	11.8	9.5	b	11.7	9.0	b	b	b
5. 2,250-5,000	35.9	14.9	b	11.5	a	a	a	a	a	a	a
6. 5,000-10,000	b	12.7	b	b	a	a	a	a	a	a	a
7. 10,000 and over	b	11.1	b	b	a	a	a	a	a	a	a

NOTES: a. No centers of this level.

b. No centers of this type

National Growth Rate 1960-70 was 13.3%

Scatter diagrams confirm those tabulated results. In those areas with the greatest reliance on agricultural earnings, the growth rate reached its lowest levels (Figure 10) while increasing federal earnings were positively related to population growth (Figure 11). Surprisingly, greater manufacturing earnings contributed only greater stability to the population growth rate (Figure 12). However, the greater the growth rate of manufacturing earnings, the greater the population growth rate (Figure 13).

The Fabric of Hierarchical Growth

Now all of this so far, the numbers of illustrations notwithstanding, has been preface. As context for our study of the fabric and symmetry of growth in the U.S. urban hierarchy in the decade, it should serve to indicate that just as regional growth has changed in form, so we should be prepared to discover that the role of local growth centers in the regional development process is also changing. The rationale for the use of urban hierarchies as the framework for growth center policies should, of course, by now be as familiar to all as were the broad features of national growth in the 1960's that I have just discussed. However, let us again refresh our memories. The argument runs as follows, as articulated by John Friedmann and others:

> Cities are the instruments whereby specialized sub-regions are articulated in a national space-economy. They are the centers of activity and of innovation, focal points of the transport network, locations of superior accessibility at which firms can most easily reap scale economics and at which industrial complexes can obtain the economies of localization and urbanization. Agricultural enterprise is more efficient in the vicinity of cities. The more prosperous commercialized agricultures encircle the major cities, whereas the inaccessible peripheries of the great urban regions are characterized by backward, subsistence economic systems.
>
> Two major elements characterize this spatial organization:
>
> (a) A system of cities, arranged in a hierarchy according to the functions performed by each.
>
> (b) Corresponding areas of urban influence or urban fields surrounding each of the cities in the system.
>
> Generally, the size and functions of a city and the extent of its urban field are proportional. Each region within the national economy focuses upon a center of metropolitan rank, and it is the network of intermetropolitan connections that articulates the whole. The spatial incidence of economic growth is a function of distance

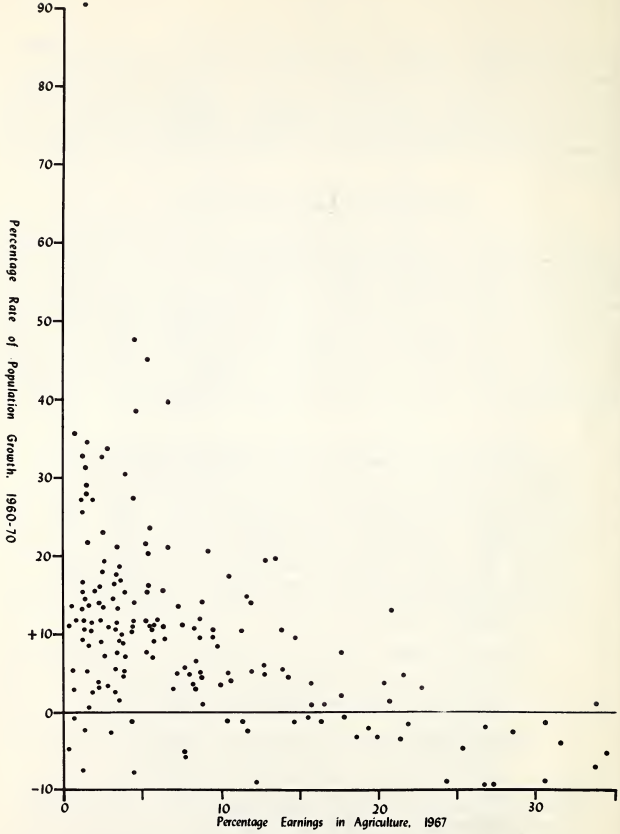


FIGURE 10 : AGRICULTURAL EARNINGS RELATED TO GROWTH RATES, 1960-1970

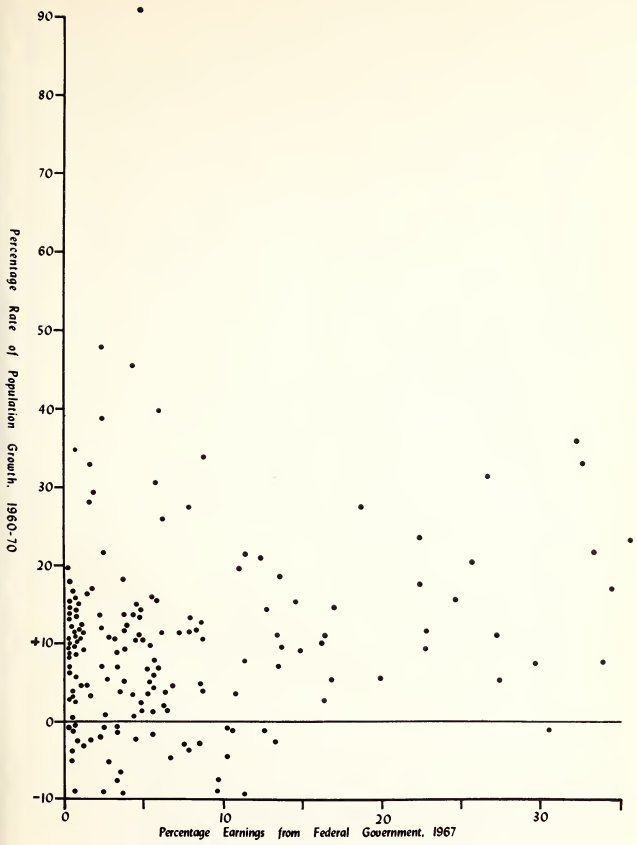
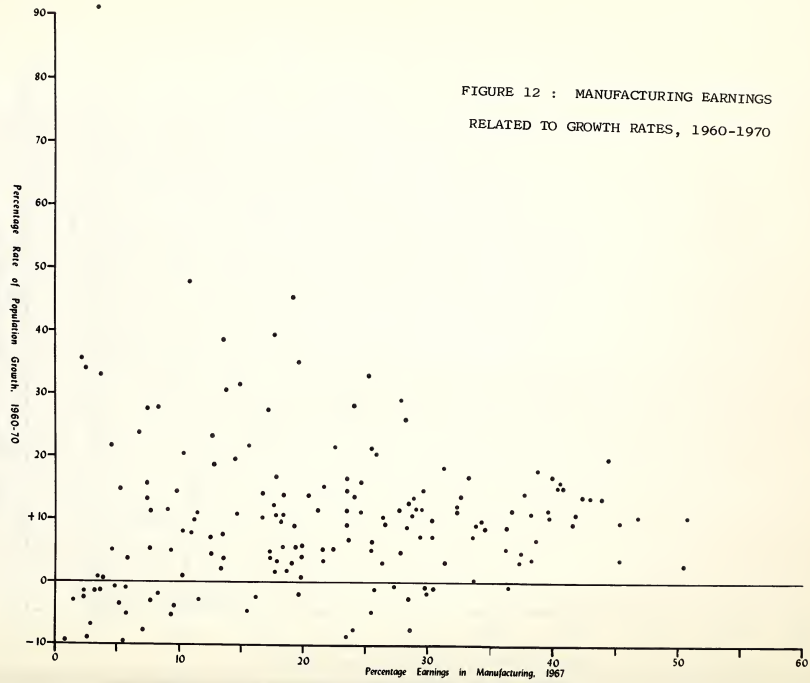


FIGURE 11 : FEDERAL EARNINGS RELATED TO GROWTH, 1960-1970

FIGURE 12 : MANUFACTURING EARNINGS RELATED TO GROWTH RATES, 1960-1970



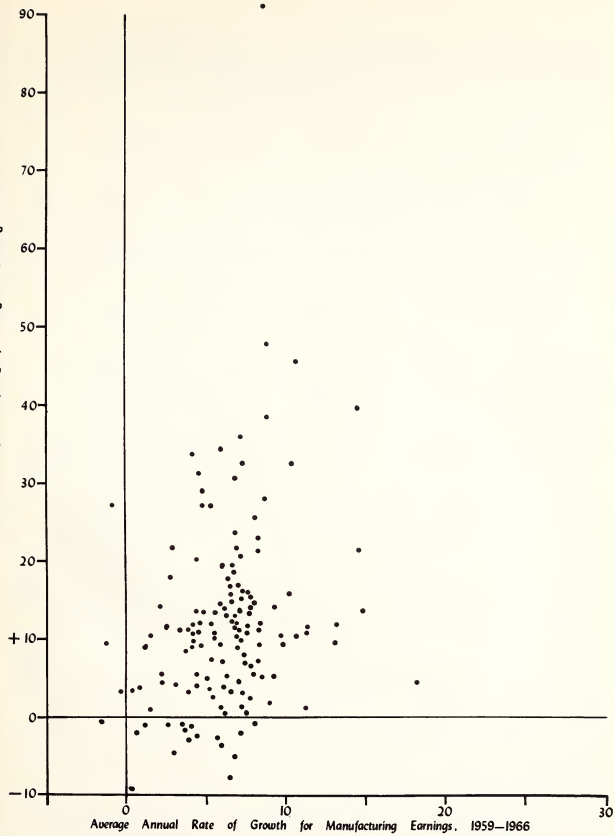


FIGURE 13 : RATE OF GROWTH OF MANUFACTURING EARNINGS RELATED TO POPULATION GROWTH RATES, 1960-1970

from the metropolis. Troughs of economic backwardness lie in the most inaccessible areas along the intermetropolitan peripheries. Further sub-regional articulation is provided by successively smaller centers at progressively lower levels of the hierarchy - smaller cities, towns, villages, etc.

Impulses of economic change are transmitted in such a system simultaneously along three planes:

(a) Outward from heartland metropoli to those of the regional hinterlands.

(b) From centers of higher to centers of lower level in the hierarchy, in a pattern of "hierarchical diffusion."

(c) Outward from urban centers into their surrounding urban fields.

Part of the diffusion mechanism is to be found in the operation of urban labor markets. When growth is sustained over long periods, regional income inequality, for example, should be reduced because the higher the capital-labor ratio in a region, the higher the employment level of the unskilled at any wage rate and at any given social minimum, therefore, the smaller the number of involuntary unemployed. Any general expansion in a high-income area, such as a heartland metropolis, will reach a rising floor to the wage-rate first. Some industries will be priced out of the high-income labor market and there will be a shift of that industry to low-income regions, i.e. to smaller urban or more peripheral areas. The significance of this "filtering" or "trickle-down" process lies not only in its direct but also in its indirect effects. If the boom originated in the high-income region, as is very likely, the multiplier effects will be larger in the initiating region, although the relative rise in income may be greater in the underdeveloped region. But the induced effects on real income and employment may be considerably greater in the low-income region if prices there are likely to rise less and/or if the increase in output per worker could be greater. Both are likely, because of decreasing cost due to external economies stemming from urbanization of the labor force. If the boom can be maintained, industries of higher labor productivity will shift units into lower-income areas, and the low-wage industries will be forced to move into even smaller towns and more isolated areas.

From this comes the basic growth center idea, and the required empirical framework -- the identification of the hierarchy, the areas and gradients of urban influence, and the diffusion times attending processes of filtering and spread.

Within such a framework, Gordon Cameron describes the growth center idea as involving spatial selectivity within declining peripheral regions into which diffusion processes have failed to carry growth. This selectivity involves official encouragement of spatially unbalanced growth within a region in the short run, by concentrating it in centers displaying the desired configuration of areas of influence, in the attempt to simulate the growth characteristics of prosperous centers in fast-growing regions with the overall objectives of attracting exogenous capital, of encouraging local product and production innovation, and ultimately of spreading economic improvements into the non-growth center parts of the region. This combination of major public investment, it is claimed, should result in an alleviation of regional balance of payments disequilibrium and should also help in substituting intraregional for interregional movements of capital and labor that threaten to denude the lagging region of critical resources. In short, then, several arguments which a priori have seeming validity, are advanced in favor or growth center strategies:

(1) such a policy contributes in the long-run, it is held, to a more rapid concentration of a region's population into a few large urban areas, with the result of creating the conditions for servicing net and replacement demand for social/economic overhead capital at low cost;

(2) for a given subsidy cost, spatially concentrated investment is likely to maximize the flow of income to regional earners in the short-run; to attract a maximum flow of exogenous enterprise and capital; and to generate a highly productive environment in which an expanded export base can reduce the regional balance-of-payments deficit and provide sufficient job opportunities to restrain the flow out of the region of economically active;

(3) the quality of short-term regional planning may be improved if the mix and scheduling of public investment over time is given a rigorous spatial dimension.

An important point should be emphasized here. Growth center "identification" does not involve the simple administrative specification of small urban centers in outlying regions as "the" growth centers for those regions. Urban hierarchies and the paths of regional interdependence are more complex than that; spatial selectivity should be accordingly more subtle.

To illustrate the problem of selection by definition, it is the position of many elected representatives and public officials in the United States that the present pattern of increasing concentration of the nation's population in major metropolitan areas is both unhealthy and undesirable. The Commission on Population Growth's surveys of residential preferences showed that a majority of Americans prefer lower-density small-town life styles. Responding to these preferences, a variety of proposals are currently be-

fore Congress characterized in common by the desire to reverse current migration flows and patterns of population change by stimulating the growth of smaller urban places in nonmetropolitan regions, i.e. by growth center strategies. Statistical support is usually provided by representatives of the U.S. Department of Agriculture. For example, representatives of the Economic Research Service have made a series of presentations before House and Senate Committees arguing that in the 1960's about 200 nonmetropolitan towns of 10,000 to 50,000 population grew by 15 percent or more, i.e. at a rate well above the national average, implying substantial "nonmetropolitan" viability of smaller population concentrations. Viability in this sense, implies economic vitality originating from within or near the identified centers, as distinguished from "mere" population growth due to an influx of a commuting population dependent on a metropolitan center. It also implies that the jobs being created are of sufficient quality to maintain, and raise in most cases, present nonmetropolitan family incomes.

Alas, when we looked at USDA's list of places, however, we found that although at least 200 nonmetropolitan towns of sizes 10,000-50,000 gained population in the 1960-1970 decade at a rate exceeding the national average according to numbers of inhabitants recorded in the 1970 census, significant portions of such increases were statistically suspect, due to annexation rather than to new growth. Many lay within the commuting areas of major cities, and grew as part of the metropolitan labor and housing market. In about 56 percent of the towns, in-migration consisted of large portions of either: 1) university or college students, 2) military personnel, 3) an elderly non-job oriented population of retirees, or 4) other institutionalized population, implying that population increase in such cases is not commensurate with growth of the labor force. Filtering out all such considerations, only seven out of USDA's list of 200 remained as "real" growth centers. Some relaxation of the growth rate criterion produced another 40-odd places. Few of these places were located in regions in which the Economic Development Administration has any continuing concern. Thus, to reiterate the point I wanted to make: identification of potential regional growth centers is not simply a matter of finding urban places in particular size ranges with particular historic growth rates. Nor is it simply the task of defining a hierarchy of "cities." We have to face the facts of the emergence of a system of interdependent and interpenetrating urban regions, with varying types and degrees of urban influence. Together, these elements provide the "symmetry, fabric and proportion" within which the relative influence on local growth of particular urban centers, operating through their local field of influence, can be evaluated, and the potential area within which spread effects are likely to be felt can be estimated.

So now to the details of our regional growth center research. First, we have defined, on a consistent national basis, the U.S. urban hierarchy. Second, within each of OBE's economic areas, we are studying the gradients of urban influence of the centers occupying each echelon of the hierarchy. Thirdly, we are examining the relationships between local job opportunities incomes, migration and population change, and relative location within the

spatial field of forces created by the urban hierarchy and the interpenetrating gradients of urban influence. The importance of the gradient analysis bears emphasizing. We seek to identify exactly which centers under what circumstances are able to command a definable hinterland that can be differentialed from the gradients of influence cast over the area by the larger metropolitan system.

The steps in the process have been as follows:

(a) All urban centers in the U.S. with populations exceeding 10,000 in both 1960 and 1970 were subjected to a 97-variable multiple factor analysis, the first factor of which indexed the relative economic "power" of each center. This analysis is described in the City Classification Handbook.

(b) On the basis of the factor scores describing economic power, the initial set of centers with populations exceeding 10,000 was stratified into the upper echelons of a hierarchy with the following levels: metropolis; major wholesale-retail center; complete shopping goods center; partial shopping goods center; complete convenience goods center.

(c) Consistency checks were made against each parallel study of regional hierarchies completed in the last two decades, including Borchert's major investigations in the Upper Midwest and my own comparative studies in several regions. These checks enabled the hierarchical stratification to be extended to all smaller urban places in the country with populations exceeding 1,500 (excluded from the multiple factor analysis because relevant census data were unavailable on all 97 input variables).

(d) The national hierarchy was reorganized by OBE economic areas, as in Table 2. All separate administrations within the core SMSA's were grouped because of their extremely close daily interdependence, and distinctions were also made to distinguish those places falling within the metropole's labor market but outside the SMSA from those places within the OBE economic area, but lying beyond its metropolitan labor market.

(e) Commuting criteria (still only available from the 1960 census) and newspaper circulation data (from the Audit Bureau of Circulations) were then used to depict the areas and gradients of urban influence.

TABLE 2

URBAN HIERARCHY: MEMPHIS, TENN. (MISS.-ARK.)
ECONOMIC AREA (OBE 046)

(a) Within Memphis Labor Market

Level	City	County	Population 1960	Population 1970	Percentage Change
Metropolis	Memphis (City)		497,524	623,530	25.3
	Shelby County		627,019	722,014	15.2
	Memphis SMSA		674,583	770,120	14.2
Wholesale-retail	—				
Complete shopping	Blytheville	Mississippi	20,797	24,752	19.0
	Clarksdale	Coahoma	21,105	21,673	2.7
	Helena	Phillips	11,500	10,415	-9.4
	Forrest City	St. Francis	10,544	12,521	18.8
	West Helena	Phillips	8,385	11,007	31.3
	Osceola	Mississippi	6,189	7,204	16.4
	Holly Springs	Marshall	5,621	5,728	1.9
	Brownsville	Haywood	5,424	6,901	27.2
	Covington	Tipton	5,298	5,771	8.9
	Oxford	Lafayette	5,283	13,846	162.1
	Marianna	Lee	5,134	6,196	20.7
Partial shopping	Wynne	Cross	4,922	6,696	36.0
	Trumann	Poinsett	4,511	5,938	31.6
	Ripley	Lauderdale	3,782	4,637	22.6
	University (U)	Lafayette	3,594		
	Bolivar	Hardeman	3,338	6,578	97.1
	Batesville	Panola	3,284	3,796	15.6
	Senatobia	Tate	3,259	4,247	30.3
	Marked Tree	Poinsett	3,216	3,208	-0.2
	Marks	Quitman	2,572	2,609	1.4
Full convenience	Sardis	Panola	2,098	2,391	14.0
	Hughes	St. Francis	1,960	1,872	-4.5
	Hernando	DeSoto	1,898	2,499	31.7
	Halls	Lauderdale	1,890	2,281	20.7
	Somerville	Fayette	1,820	1,779	-2.3
	Manila	Mississippi	1,753	1,961	11.9
	Lepanto	Poinsett	1,585	1,846	16.5
	Leachville	Mississippi	1,507	1,582	5.0

(b) Beyond Memphis Labor Market

Wholesale-retail	Jackson	Madison	34,376	39,262	14.2
Complete shopping	Jonesboro	Craighead	21,418	27,050	26.3
	Dyersburg	Dyer	12,499	13,942	11.5
	Paragould	Greene	9,947	10,639	7.0

TABLE 2 (Cont.)

Level	City	County	Population 1960	Population 1970	Percentage Change
	Paris	Henry	9,325	9,803	5.1
	Humboldt	Gibson	8,482	10,000	17.9
	Milan	Gibson	5,208	7,115	36.6
Partial shopping	Martin	Weakley	4,750	7,738	62.9
	Trenton	Gibson	4,225	4,181	-1.1
	Lexington	Henderson	3,943	4,986	26.5
	McKenzie	Carroll, Weakley	3,780	4,813	27.3
	Pocahontas	Randolph	3,665	4,544	24.0
	Walnut Ridge	Lawrence	3,547	3,800	7.1
	Bemis (U)	Madison	3,127		-
	Piggott	Clay	2,776	3,087	11.2
	Henderson	Chester	2,691	3,512	30.5
	Ripley	Tippah	2,668	3,482	30.5
Full convenience	Corning	Clay	2,192	2,705	23.4
	Huntingdon	Carroll	2,119	3,525	66.4
	Dyer	Gibson	1,909	2,430	27.3
	Hoxie	Lawrence	1,886	2,265	20.1
	Parsons	Decatur	1,859	2,119	14.0
	Greenfield	Weakley	1,779	2,051	15.3
	Rector	Clay	1,757	1,990	13.3
	Newbern	Dyer	1,695	2,081	22.8
	Marvell	Phillips	1,690	1,980	17.2
	Alamo	Crockett	1,665	2,385	43.2
	Dresden	Weakley	1,510	1,933	28.0

(f) The gradient-like patterns of 1960-70 migration rates and population change are now being added to the scheme, as will such features as income changes and changes in the distribution of the poverty population.

These materials enable us to describe the fabric and texture of regional growth in the 1960s, and to identify the existing network of local points of relative growth, and their areas of influence in a consistent manner. To return to the Memphis example, Figure 14 shows the labor market of the metropolis and also identifies those counties added to this labor market by OBE when they defined the Memphis economic area. The larger area corresponds roughly to that territory within which Memphis captures the major market share of regional newspaper sales (Figure 15). Table 2 tabulates the growth rates of urban centers within the various parts of the Memphis economic area, revealing many small places with impressive growth rates.

Figure 16 shows three traverses radiating from Memphis to the centers of adjacent economic areas. If cross sections through the map of gradients of urban influence or of migration and population change are taken along these traverses and graphed, Figures 17-22 result.

In every case the "topography" of spatially-varying migration rates and population changes among both blacks and whites rises and falls in rhythm with the gradients of urban influence of the larger metropoles and the lower-order wholesale-retail centers - compare Figures 17 and 18, 19 and 20, 21 and 22. In each case the "peaks" correspond to major urban nodes and the "valleys" to the interurban peripheries; a few of the peaks display both population growth and net immigration, most parts of this region display net outflows of people, however - more rapid for blacks than whites. The height of the peaks corresponds to the level of centers in the hierarchy. Very few of the centers in Table 2 located on the traverses show any capability of developing any spread effect around them. Within the spatially-extending labor markets of the metropoles, internal polarizing shifts are seen to be taking place. For example, the inverse black-white patterns are quite clear in the Nashville labor market in Figure 18, and also represented in the southward and westward thrusts of the Memphian white population indicated in Figures 20 and 22. Note other features, too: the westward thrust of white immigration and population increase in the outer extremes of the Nashville labor market, towards Jackson, in the only zone with too few blacks for the census to report data for blacks, and the absence of any real gradient of influence around Jackson itself (Figure 18); the subsidiary peaks of Jonesboro and Tupelo, growing in population and almost holding net migration among whites to zero (a distinct change from the 1950-60 decade), and the conformity of gradients of change around these places to the configuration of their newspaper circulation territories (Figure 20); and the suggestion that there are a few lower-order outlying poles of relative importance in generating local spread effects (Grenada and Dyersburg) in Figure 22.

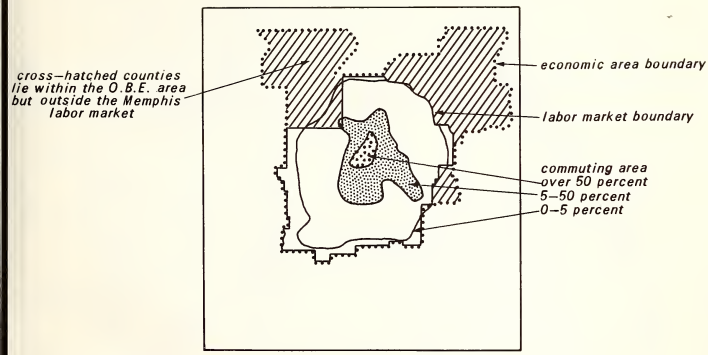


FIGURE 14: THE MEMPHIS REGION

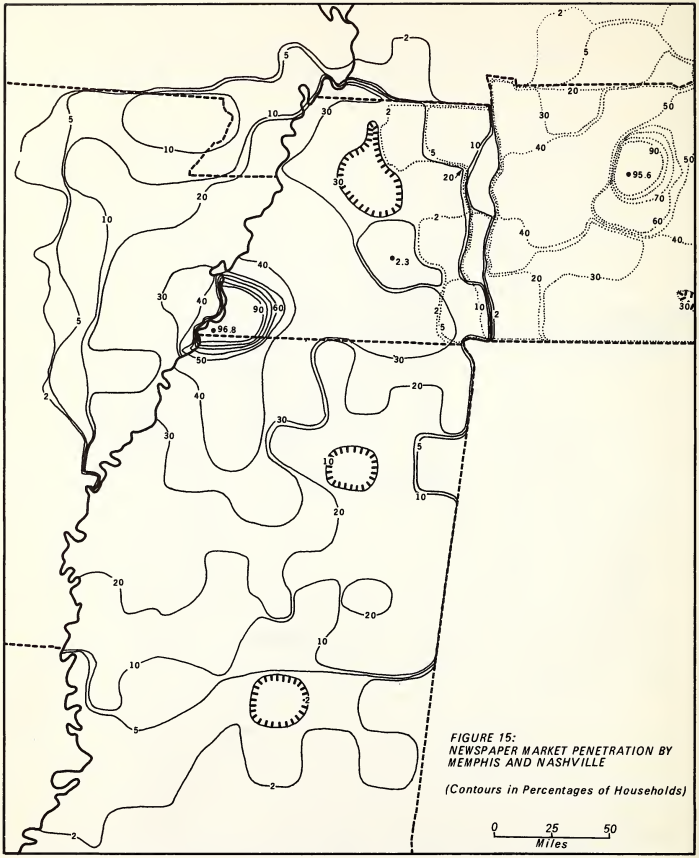


FIGURE 15:
NEWSPAPER MARKET PENETRATION BY MEMPHIS AND NASHVILLE

(Contours in Percentages of Households)

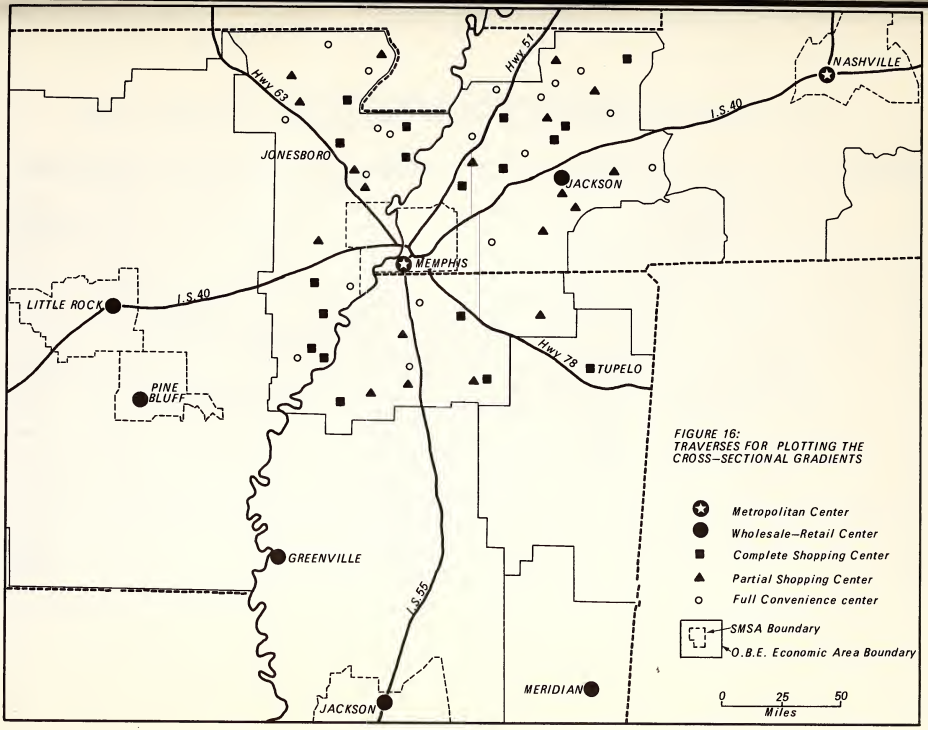


FIGURE 16:
TRAVERSES FOR PLOTTING THE CROSS–SECTIONAL GRADIENTS

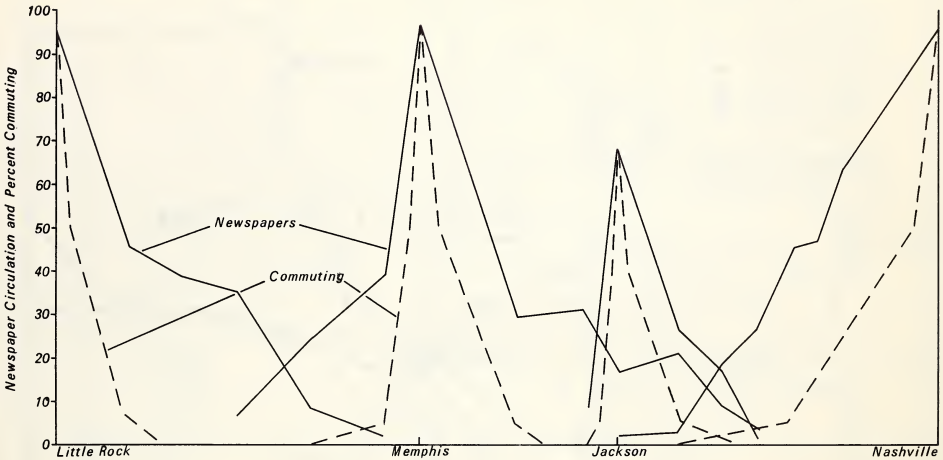


FIGURE 17: GRADIENTS OF COMMUTING (1960 DATA) AND NEWSPAPER MARKET PENETRATION (1971 DATA) ALONG ROUTE FROM LITTLE ROCK TO NASHVILLE

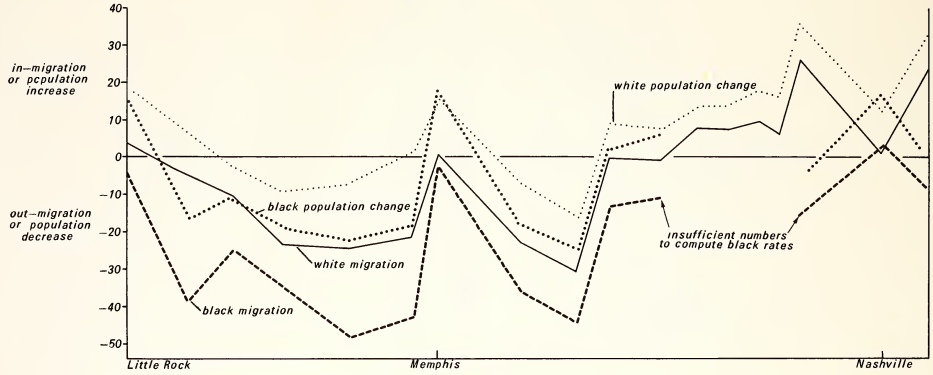


FIGURE 18: MIGRATION AND POPULATION CHANGE GRADIENTS, 1960–1970, BLACK AND WHITE

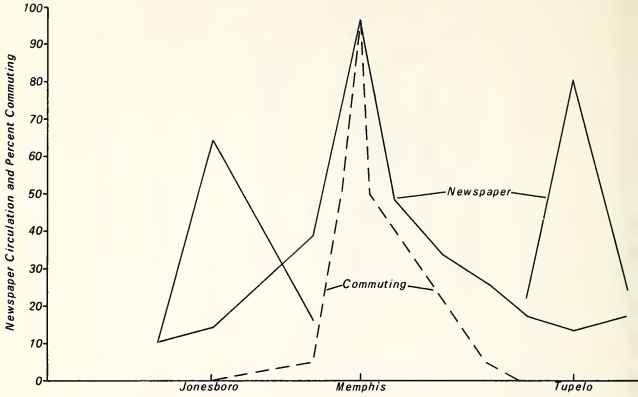


FIGURE 19: NEWSPAPER CIRCULATION AND COMMUTING AREAS, JONESBORO TO TUPELO

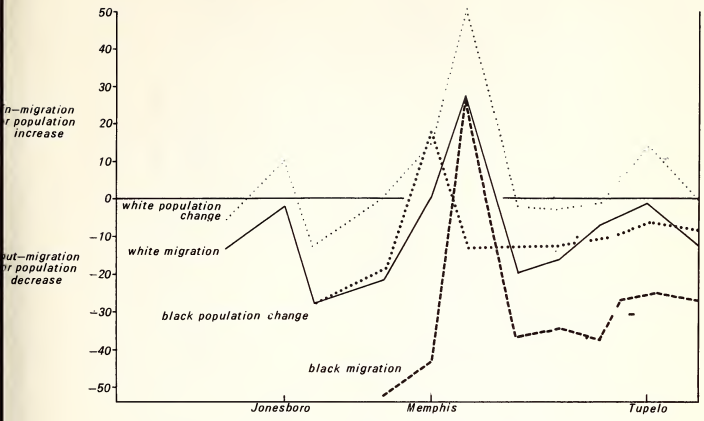


FIGURE 20: MIGRATION AND POPULATION CHANGE, JONESBORO TO TUPELO

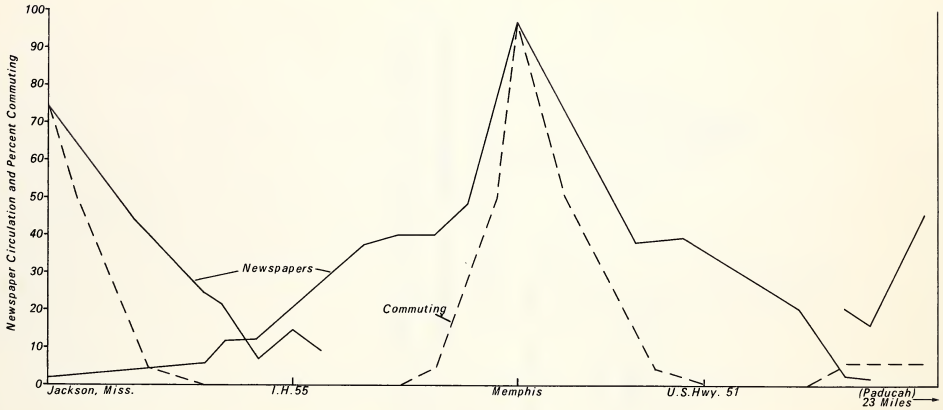


FIGURE 21: NEWSPAPER CIRCULATION AND COMMUTING FIELDS, JACKSON, MISS. TOWARDS PADUCAH

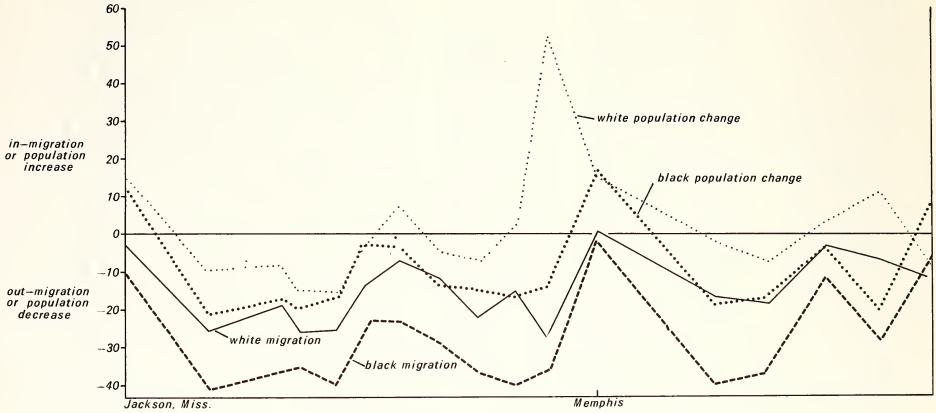


FIGURE 22: POPULATION CHANGE AND MIGRATION JACKSON, MISS. – MEMPHIS – PADUCAH

When I first agreed to write this paper, I thought that I would discuss other examples too, and I selected the Detroit economic area as one example and the Four Corners region as the other. You can already guess the results, however, so I will not inundate you with yet other illustrations. The Detroit region looks like Nashville in Figure 18, enlarged to a 50-60 mile radius (refer back to Figure 5). In Four Corners downward gradients dip into the heart of the region from Denver, Albuquerque, Phoenix and Salt Lake.

Those of you who recall the paper entitled "Spatial Organization and Levels of Welfare" that I read to an earlier EDA research conference will recognize the repetitive rhythms that I am talking about. Their essential nature does not appear to have changed in any major way between 1960 and 1970; the cumulative incapacities that accompany lower status and greater peripherality persist. Growth differentials are patterned by relative accessibility to job opportunities in a hierarchy of metropolitan and urban regions; the opportunities decline as one proceeds down the hierarchy. And the population is responding to the differentials in opportunity, blacks relatively more rapidly than whites.

The immediate impression for those who were committed to the attempt to change this in the 1960's might be that we are describing a complex system with all the perversity associated with such systems by Jay Forrester, inherently insensitive to policy changes and drawing our attention to "problem areas" which are the very points at which attempts to intervene are most likely to fail because they represent the ends of processes operating in society rather than the beginnings.

But all the evidence is not yet in. To the kind of "fabric" picture that we have just described for Memphis - and will provide for all 171 continental OBE economic areas later this year - we have to add materials on changes in the nature of change (i.e. in migration rates, population change etc.) 1950-60:1960-70, as well as changes in income levels, in farmland values, in commuting patterns, and in accessibility, and their associations. We cannot do this until the fourth count census data are delivered, of course, but we are ready to go once the census gives birth. Why should this final step be important? It is because I believe that enough suggestive nationwide evidence has already begun to accumulate, for example in cases like Jonesboro and Tupelo, to suggest that, the broad continuities of pattern notwithstanding, the behavior or the system and the subtle details of its local patterns are changing and being changed in ways that reflect actions taken or new processes initiated at senstive influence points, local growth centers capable of exerting limited but independent spread effects. Most of these influence points tend not to be self-evident; gradient analysis is required to distinguish them in their local setting. Yet spatially-selective policy undertaken at them may indeed produce pressures that radiate to other parts of the system in desired ways. Only a careful examination of local dynamics can reveal such points and the areas into which radiating

influences flow. What I hope is that by the conclusion of our research that is exactly what will have been provided: a knowledge of the specific points that mediate the relationships between urban growth and regional change, so that we can put our fingers on those points in space at which provision of new opportunities is most likely to call forth responses in the problem regions that are of continuing concern.

BEHAVIORAL MODELS OF INTER-REGIONAL MIGRATION

IRVING N. FISHER

I. INTRODUCTION

Recent programs of EDA and other Federal agencies have focused on industrialization of rural areas having limited economic potential -- areas characterized by chronic unemployment and a narrowing economic base. Most of these programs, however, have had only limited success in stimulating development in these regions.

It is commonly believed that out-migration has been one of the major factors limiting the success of these programs. Critics argue that migration to larger cities and towns offering more attractive employment opportunities siphons off the younger, better educated segments of the population, leaving a residual that is lacking in many of the skills and attributes necessary to provide a productive and efficient labor force. Consequently, advocates of rural development programs have emphasized the need to halt and, indeed, reverse the flow of migrants to larger metropolitan areas.

Although migration would appear to present a major obstacle to rural economic development, there is insufficient information about the magnitude and patterns of rural-urban, South-North migration to determine precisely its impact on the development potential of these rural areas. And, although the literature is extensive, specific knowledge of the nature and causes of rural-urban migration is grossly inadequate for the task of predicting the impact of population movement on rural and regional growth policies.

This paper first examines the present status of our knowledge of the inter-regional migration process and then reports on a current study which might increase that knowledge.

II. LIMITATIONS OF EXISTING MODELS OF INTER-REGIONAL MIGRATION

Most studies of migration provide only a limited understanding of the mechanisms underlying the migration process. Perhaps the most serious shortcoming is the lack of a realistic theoretical framework for explaining and analyzing observed patterns of population movement. While a number of investigators have been successful in relating observed patterns of migration to various socio-economic variables, these studies have limited value either as tools for predicting future population movement or for providing a framework for evaluating alternative policy options.

The major difficulty stems from the degree of generalization and over-simplification characteristic of most theoretical treatments of the migration process. Present theories often fail to consider the numerous non-economic

factors that influence the decision to migrate and the choice of destination. They also ignore differences in both the observed rates and patterns of migration among major subgroups of the population, and fail to explain the striking differences between inter-metropolitan and rural-urban migration patterns. None of the existing theoretical models of migration explain the linkage that appears to exist among subsequent moves, especially for rural-urban migrants.

Much of the research directed at developing a theoretical framework for understanding and explaining population mobility has focused on two different aspects of the migration process: characteristics of the geographical location and characteristics of the migrant. At one extreme are a variety of studies that identify and attempt to explain aggregate patterns of population movement between various origins and destinations. Other studies, in contrast, focus on identifying the motivational forces that lie behind the individual's decision to migrate.

Aggregate Models of Population Mobility

Most aggregative population models attempt to explain observed patterns of population movement between origin and destination points as a function of differences in various attributes and characteristics of the origin and destination areas. In short, aggregative-type migration models attempt to allocate or redistribute population among competing geographical regions as a function of the relative attractiveness (or unattractiveness) of the regions.

Virtually all of the aggregative-type migration models are variants of the well-known "gravity" models, which take the form:

$$M_{ij} = \frac{K\, P_i P_j}{D_{ij}}$$

where M_{ij} = net/gross migration between regions i and j;

P_j = population in region j;

D_{ij} = distance between region i and j; and

K = constant term.

A number of recent studies of inter-regional population movement have utilized the basic gravity model to explain observed patterns of migration between selected origins and destinations.* These studies typically replace the constant term with a set of variables indicating differences in such

*Cicely Blanco, "The Determinants of Interstate Population Movements," Journal of Regional Science, V, 1 (Summer 1963), 77-84, and "The Determinants

factors as income, wage rates, unemployment rates, and racial balance. Lowry has perhaps extended this theory as far as possible by including variables to denote unemployment, wage rate and other demographic characteristics in both origin and destination locations.*

A major difficulty with these gravity-type models stems from the level of aggregation utilized in estimating the parameters of the equations. There are indications that the mechanics of the migration process differ significantly for various sub-components of the migrant population. Migration of low-income blacks, for example, seems to be independent of relative labor market variables, but closely related to such factors as ethnic similarity of origins and destinations. White migrants in the middle- and upper-income groups, however, move more often as a result of job changes or in response to economic factors. Consequently, further research focusing on differential rates and causes of migration among various subgroups of the population could provide a basis for constructing a more realistic theory of migration and, more specifically, provide additional insight into the rural-urban migration process.

A second major difficulty with these models results from the extremely limited number of explanatory factors that are utilized to explain observed patterns of migration. Not only is it quite likely that migration patterns differ among various subgroups in the population, but also that these groups are influenced by different motivating forces. It is apparent, for example, that some ethnic groups move to places where friends or relatives already reside. In contrast, older persons typically move in response to climate and geography while others may relocate because of job transfers. In other words, it is probable that other noneconomic factors play an important role in determining patterns of population movement, but that these effects are hidden by the aggregation of the data.

Perhaps the most bothersome difficulty with these aggregative migration models concerns the conclusions and inferences that have been based on the

---

of Regional Factor Mobility," Ph.D. Diss., Nederlandsche Economische Hoogeschool te Rotterdam, 1962; R. McGinnis et al., "Internal Migration as a Stochastic Process," Bulletin de l'Institut International de Statistique, Toronto, XL, 1 (1964), 446-47; R. L. Mills, "A Transition Matrix Approach to Migration Analysis and Population Forecasting," Ph.D. Diss., Indiana University, 1967; Robert L. Raimon, "Interstate Migration and Wage Theory," Review of Economics and Statistics, XLIV, 4 (November 1962), 428-38; T. W. Rogers, Differential Net Migration Patterns in Southern SMSA's, 1950-60," Ph.D. Diss., Mississippi State University, 1966; Samuel A. Stouffer, "Intervening Opportunities and Competing Migrants," Journal of Regional Science, II, 1 (1960), 1-26; and J. D. Tarver et al., "A Stochastic Analysis of Geographic Mobility and Population Projections of the Census Divisions in the U.S.," Demography, II (1965), 134-39.

*I. Lowry, Migration and Metropolitan Growth: Two Analytical Models, San Francisco, Chandler, 1966.

findings of these studies. The statistical results that have been obtained suggest that conditions in the place of origin do not play an important role in explaining observed patterns of migration. This has led a number of researchers to propose that migration be characterized as a two-stage decision process in which the decision whether to move is made independently from the decision where to migrate. The problem, of course, is that separating the two components of the migration process masks the obvious connection that exists between the decision to move and the choice of destination. While it is undoubtedly quite useful to identify those features and attributes that increase one's propensity to migrate, it is, nonetheless, essential to retain the connection between the two decisions in the theoretical framework. The reason should be obvious: in some cases the propensity to move bears a direct relationship to the availability of an attractive and suitable destination. A case in point would be the alleged connection between the relatively high welfare benefits in northern cities and the movement of poor black families into those cities from the South.

The Individual's Decision to Migrate

As noted above, considerable literature concerning the decision to move has appeared in recent years, resulting in a real dichotomy between those studies that explain aggregate patterns of population movement and those that distinguish migrants from non-migrants.* Much of the work that has been done in this latter area in the past decade has been partially motivated by the inability of the broader-brush, aggregative models to fully account for the "why's" of migration. More particularly, these micro-analytical studies have focused on the "why not's"; they explore why people who are faced with an apparently economically-appealing choice of location choose not to migrate there.

---

*K. C. Land, "Duration of Residence and Prospective Migration: Further Evidence," Demography, VI, 2 (May 1969), 133-40; J. B. Lansing and W. Ladd, The Propensity to Move, Labor Mobility Publication 3, U.S. Area Redevelopment Administration, Washington, D.C., Government Printing Office, 1964; Everett S. Lee, "A Theory of Migration," Demography, III, I (1966), 47-57; P. A. Morrison, "Duration of Residence and Prospective Migration: The Evaluation of a Stochastic Model," Demography, IV, 2 (1967), 553-61; H. S. Shryock, Jr., Population Mobility Within the United States, Community and Family Study Center, Chicago, University of Chicago Press, 1964; H. S. Shryock, Jr., and E. A. Lamon, "Some Longitudinal Data on Internal Migration," Demography, II (1965), 579-92; K. E. Taeuber, "Duration of Residence Analysis of Internal Migration in the U.S.," Milbank Memorial Fund Quarterly, XXXIX, 1 (January 1961), 116-31; G. L. Wilber, "A Bayesian Model for Migration Decisions in a Population," paper presented at the Annual Meeting, Population Association of America, Chicago, 1965; and J. Wolpert, "Behavioral Aspects of the Decision to Migrate," Papers and Proceedings of the Regional Science Association, XV (1965), 159-69.

One attempt to express the migration decision within the social scientist's classical utility maximization framework was suggested by Wolpert.* However, as Morrison later objected in reference to the utility approach, "There is little prospect that the underlying utility function or its calculus could be empirically ascertained with any degree of precision.** Morrision, therefore, chooses to expand upon Lee's suggestion*** that the migration decision is a two-part phenomenon; first, a decision to move and, second, a choice of location. In more recent work, Morrison concentrates on the process of deciding to move.**** In these papers Morrison hypothesizes the existence of a migration threshold that varies among individuals depending on age, race, education, income, family status, and duration of residence in a particular region. It is suggested that the continuous variation of these characteristics among the entire population could be aggregated into a more manageable number of discrete elements in matrix form.

Although considerably more useful in identifying those differences that distinguish migrants from non-migrants, the analysis ignores any connection that may exist between the decision whether to move and the choice of destination. As a result, it is difficult to analyze the linkage between demographic characteristics of various components of the migrant population and observed patterns of migration. Thus, more elaborate and comprehensive models are needed to develop the link between the two decisions and provide a framework for analyzing the efficacy of various policies and programs that affect the decision to move.

Inadequacy of Empirical Data

One of the major factors limiting our understanding of the migration process has been the lack of a suitable data base upon which more realistic migration models could be based. Most empirical evidence presently available is derived from cross-sectional data collected as part of the decennial census. Other migration statistics are typically estimated as a residual component of population change. Consequently, data identifying the frequency and sequence of moves is virtually non-existent and, as a result, it has been impossible to construct a comprehensive time series for population movement in the U. S. What is needed is a longitudinal time series providing a continuous history of individual migration patterns.

---

*Wolpert, op. cit.

**Morrison, op. cit.

***Lee, op. cit.

****P. A. Morrison, "Implications of Migration Histories for Model Design," paper presented at the Conference on Migration Histories, sponsored by the Center for Population Research, National Institute of Child Health and Human Development, Washington, D.C., February 5-6, 1970, and "Theoretical Issues in the Design of Population Mobility Models," paper presented at the 64th Annual Meeting, American Sociological Association, San Francisco, September 1-4, 1969.

III. A CURRENT STUDY OF THE RURAL-URBAN MIGRATION PROCESS

Research currently underway examines some of these issues and provides insight into the mechanics of the rural-urban migration process. Funded by the Office of Economic Research, the study has had two principal objectives: to identify the linkage and sequence of moves in the rural-urban migration process, and to identify the impact of this migration on the labor supply and population of rural areas.

Rural-Urban Migration Linkage

It is often assumed that migrants leaving rural areas in the South and Midwest move directly to larger centers of the North and West. Recent evidence, however, suggests that many of these migrants previously lived in smaller metropolitan centers prior to migrating to major urban areas. Consequently, the rural-urban migration process may be more complex than thought involving a sequence of moves to progressively larger urban areas, perhaps initially within the same geographic region but eventually involving inter-regional moves.

This study utilizes two sources of data: a ten year continuous sample from the Social Security Administration's Continuous Work History File and the 1966 and 1967 Survey of Economic Opportunity. The former identifies the county of employment for individuals migrating from rural areas of the South to the North over the ten year period of 1957-1966. It is thus possible to establish the linkage -- the sequence and stages of moves -- between rural areas of the South and more highly industrialized areas both in the South and in the North. In addition, the data indicate changes in employment and income which result from various types and patterns of moves and provide a basis for identifying the policy implications for regional growth center programs and policies. Some of the more interesting conclusions to emerge are the following.

1. Patterns of migration from the rural South. Although there has been considerable speculation that the rate of migration from the rural South may have declined during the 1960's, the evidence suggests that the movement of population from rural areas of the South and midwest to larger metropolitan regions in the North continued through the ten year period at a rate very nearly equal to that observed in earlier periods. In addition, the evidence indicates that:

(a) Mobility among residents of the rural South seems relatively high. Evidence indicates that more than 70 percent of those employed moved to a different county at least once during the ten year period.

(b) A majority of this group, however, moved to nearby or contiguous counties; moves between more distant counties occurred less frequently and declined markedly with increasing distance of move.

(c) Nonetheless, more than one-fourth of those who moved during the period made at least one move of more than 1,000 miles.

(d) As would be expected, there are an enormously large number of unique patterns of movement. Aggregation into roughly comparable groups still results in a significant number of distinctly different patterns -- some 40 to 50 different groups. Some of the more interesting patterns that emerge are:

- o a large number of moves take place in a repetitive fashion between the same set of counties located within a rather narrow geographical area;
- o there are a significant number of round-trip moves -- moves from rural areas of the South to the North and back again -- suggesting that return migration may be more prevalent than commonly assumed;
- o a noticeable number of moves of more than 100 miles occur in apparently random fashion among nearby counties within the same general geographic region, indicating a rather high degree of intra-regional mobility;
- o a substantial number of moves also occur repetitively between predominantly rural, non-metropolitan counties in different geographic areas of the South, suggesting a rather substantial degree of inter-regional mobility among rural residents of the South.

2. Return migration from the North. Utilizing the same source of data, the study also examines the movement of people from the larger, more highly industrialized areas of the North to the South. In particular, the study identifies patterns of movement between northern metropolitan areas and rural and urban areas in the South and identifies various socioeconomic characteristics of these return migrants. In addition, the study examines the implications of return migration for various regional and rural growth center programs and policies.

Evidence suggests that there are a substantial number of migrants from northern metropolitan areas to the South Atlantic states and that this flow may have increased during the 1960's. In particular, the data indicate that:

(a) The movement of people from northern metropolitan areas to the South is surprisingly large, particularly among the black component of the population (approximately one-third of the migrants moving between northern cities and southern areas were black.)

(b) This trend appears to have increased significantly during the ten year period, especially among the black component of the population.

In addition, these return migrants are typically better educated, more highly skilled, and earn larger incomes than their rural counterparts. In fact, there is some indication that major urban centersof the North serve as training grounds, importing relatively unskilled labor from rural and urban areas of the South and exporting more highly skilled workers to these same regions. In fact, it appears that:

(c) While a majority of these return migrants do become employed in predominantly urban, metropolitan areas of the South, there are an impressive number employed in surrounding rural counties and also in predominantly rural counties in the less developed areas of the South.

(d) Although it is difficult to make generalizations, it would appear that return migration from the North may provide a valuable source of skilled labor both for the more highly urbanized areas of the South and predominantly rural areas.

IV. CONCLUSION

Although the evidence forthcoming from this study does promise considerable insight into the complexity of the rural-urban migration process, it is but a first step toward developing a more realistic understanding of the processes underlying interregional population movement. Extensive analyses of the various causal factors underlying the observed migration patterns that emerge are necessary before any generalizations or conclusions can be made. In short, what is needed is a more comprehensive dynamic model of inter-regional mobility, a model that encompasses not only the pattern and sequence of moves but also includes the effects of a broad range of both economic and non-economic factors.

NATIONAL GROWTH POLICY AND THE ENVIRONMENTAL EFFECTS OF CITIES

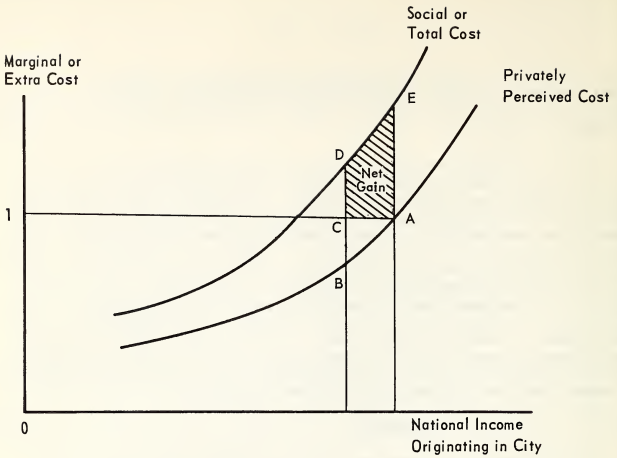
GEORGE S. TOLLEY

The present paper is a progress report to estimate the environmental effects of population distribution policies. The first section presents a framework for the analysis. The framework is then used to estimate the change in national daily travel costs resulting from a policy of retarding growth of large cities. As an example of a factor that has external effects that are not usually privately perceived, attention is next given to urban density. Gains in external density caused by increasing a smaller city's population are estimated for two public services. As further examples of factors with external effects, in depth refinements are developed for estimating congestion and pollution effects of urban growth, giving attention to the effects of road capacity decisions and to the complexities of air pollution costs within a city. The closing section of the paper discusses the broader implications for national growth policy.

Framework

The total change in income resulting from altering the location of activity is the sum of the increases in income for areas where economic activity is made greater, minus the sum of the income decreases for areas where economic activity is diminished. The invisible hand assumption is that individual actions are carried to the point where extra gains just equal extra costs. If this assumption were fulfilled, the output of each city would be carried to the point where the cost of additional units of output was equal to the value of additional output. The market economy would achieve the spatial distribution of activities making the greatest contribution to national income. Any policy altering this distribution would have the net effect of decreasing national income. For cities reduced in size by a policy, output would be diminished by more than costs. For cities whose size was augmented, output would be raised by less than their increase in costs.

If the invisible hand assumptions were fulfilled, the top and bottom parts of Figure 1 would each contain only one curve. In the top part of Figure 1, for a city reduced in size from A to C, each unit of reduction in output would reduce output by a dollar, implying output movement along the horizontal line AC. Cost reductions would consist entirely of reductions in private costs, which would be progressively reduced moving along the marginal cost curve AB. The excess of the output reduction over the cost reduction would be the area ABC. By similar reasoning in the lower part of the figure, for a city increased in size from F to H the excess of the cost increase over output increase would be the area FGH. The decrease in national income from a policy affecting the distribution of activity would be the sum of the ABC areas for cities reduced



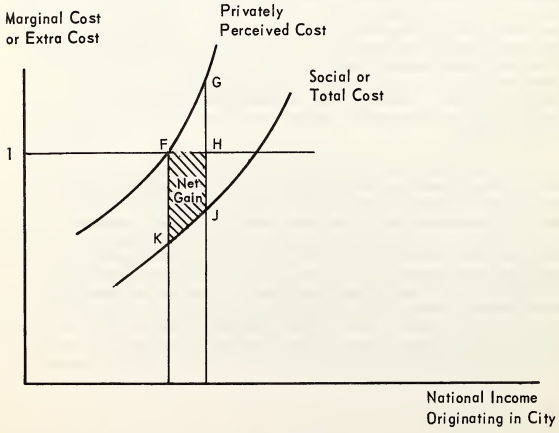


FIGURE 1. Changes in City Size

in size plus the sum of the EFG areas for cities increased in size.

There are many reasons why the invisible hand assumption is not fulfilled. In the top part of Figure 1, a situation is shown where costs are imposed on others. The costs are largely unperceived by the individuals and firms causing them and are not taken into account in their actions. Pollution and congestion are examples of these largely unperceived external costs imposed on others when there is additional output in a large city. In Figure 1, the external cost imposed on others from an additional dollar of output in the city is AE. Instead of there being a net cost to reducing the size of the city from A to C, there is a net gain. The output reduction before was given by movement along the horizontal line AC, whereas the cost reduction is now given by movement along the higher curve ED. The net gain from the reduction in city size is the difference between the private cost - area ABC and the external gain area ABDE, or the shaded area ACDE.

The effect of the policy for the city in question can be estimated as the net sum of the two parts just noted. With regard to external costs ABDE, the change in city size is unlikely to greatly affect marginal external costs AE for most realistic policies, since policies cannot realistically be expected to change sizes of large cities by more than a few percent. The change in external costs can thus be estimated as marginal external cost AE times change in output AC. The other part of the policy effect, which is the excess of the change in output over the change in private costs, is the area ABC previously considered. The effect on income ACDE from changing the size of the city is thus change in external costs ABDE minus the excess of change in output over change in private costs ABC. Estimating the area ABC requires quantifying how private costs vary with level of activity. Estimating the area ABDE requires quantifying externalities.

The lower part of Figure 1 depicts a situtation where there are external economies making costs less than privately perceived costs. Economies of scale are frequently cited as a phenomenon associated with city size. The economies can be internal or external. An example of an external economy is the effect of an increase in city size increasing density and thereby reducing costs of delivery and pickup. Pupil transportation and waste pickup services are two of many examples of functions whose cost is reduced by an increase in density. Economies due to density are external or unperceived in private decisions. People's bids for land, which determine density, reflect only the gains and costs to themselves taking as given the conditions facing them in the city. The reduction in costs of travel by other individuals, firms or government entities due to greater density is not a part of the calculation influencing the individual actor's decision. Density is a production function shifter having the technical attributes of a public good. It shares the problem encountered with such goods that the full gains from having the goods are not reflected in individual market decisions in the absence of government or other outside arrangements. Density is only one of several reasons for external economies of scale encountered at some stages of city growth. In the lower part of Figure 1, the gain from increasing city size from F to H is FHJK, estimable as external economy FGJK

less excess of additions to private costs over additions to output FGH.

Hopefully a population distribution will reduce city sizes for which there is an economic gain from city size reduction and will increase city sizes for which there is an economic gain from expansion. These are the situations depicted in Figure 1. The gains could, however, be negative if the relation between private and social costs happened to be reversed. In the upper part of the figure, the social cost curve might lie below the private cost curve, and in the lower part of the figure social might be above private cost. For unemployment effects and for institutional or public finance effects, the relation between social and private cost might vary geographically by region but not necessarily in any systematic way by city size. These effects thus might make social costs greater or less than private costs.

Distance to Work

The preceding section indicates that the estimation of effects of population distribution policy subdivides into a concern with 1) external effects and 2) changes in private or direct costs from changing the amount of economic activity in an area. Among the reasons for changes in private costs is commuting and related daily travel. In a small town, these costs are small, and they increase with increasing city size. The rise in daily travel costs is one of the limiting considerations determining city size. The cost advantages from having some kinds of production in a large city are sufficient to compensate for large daily travel costs, but the rise in costs due to daily travel eventually raises the cost of additional output in the city to that in smaller cities. Daily travel costs appear to be the main reason why most of the nation's economic activity is not located in one huge city.

Complaints of residents of cities and general indignation over environmental conditions in larger cities have given impetus to recommendations that attempts be made to foster a growth pattern away from larger cities. The recommendations could be justified on national income grounds if the reduction in external environmental costs more than offsets the loss due to the greater reduction in output than in private costs, that is, if ABDE exceeds ABC in Figure 1. When city size is reduced, or when a city is kept from growing as much as it otherwise would, it is kept away from equilibrium point A. As reflected in market costs of resources indicating their opportunity returns, the labor and capital resources excluded from the city are producing less private returns elsewhere than if they were admitted to the city.

Suppose that the growth of an SMSA of 6 million persons is stopped for a decade, during which time its population would otherwise grow by 10 percent. In Figure 1, the city is kept at C instead of being allowed to go to point A. If daily travel were the only reason for upward slope of the marginal cost curve, an estimate of the effects on the travel would enable estimation of the area ABC.

The number of persons living in a ring of width dr located r miles from the central business district is the area of the ring $2\pi r\,dr$ times the population density of the ring D_m. In this example the distances at issue are on the fringes of a large city where typical suburban densities may be expected to prevail, suggesting a reasonable assumption is a value of D_m of about 1000 persons per square mile. Let c be travel cost per mile, which will be assumed here to be \$.10 per mile. Let τ be the daily travel per person to the central business district and to other destinations not affected by where the family lives, expressed as the number of round trips that could be taken to the central business district if the travel mileage were for such trips only. If the population consisted entirely of four person households, and if the only effect on travel of living further away from the central business district were to increase the commuting trip of the head of the household to the central business district, τ would be 1/4, i.e. one out of every four people making one standard trip each day. A consideration decreasing τ is that not all work trips are to the central business district. The fact that many shopping trips and goods deliveries involve that travel destination tends to increase τ. Bearing in mind that τ is for travel from the outlying points of the city where growth takes place, refined estimates of τ could be made from travel surveys. In this example, it will be assumed that τ does not change over the range of city growth being considered. This assumption is probably defensible for the assumed growth range of only 10 percent. For a greater percentage range of growth, particularly if one were considering growth of a smaller city, τ would be expected to decline as distance from the central business district increased. This effect could be allowed for by expressing $\tau = \tau(r)$. The modifications of the formulas below are obvious and could be handled in exact numerical analysis.

Since the mileage for a round trip to the central business district is twice the distance to the central business district, there is extra mileage 2m-2r of travel from the edge of the city over travel from a ring at distance r. Let m_1 refer to the distance to the edge of the city that would prevail if the city reached private equilibrium. The city size corresponding to m_1 is at point A in Figure 1. The mileage $2m_1$-2r is extra travel it would be worthwhile to undertake from closer distances r and still cover the costs of output produced by persons living at r. For distances to which the city is not allowed to grow, the cost of travelling $2m_1$-2r miles is a measure of the net gain foregone from not having people live at distance r. The sum of extra costs from the actual margin m_o, to which the city is constrained by a population distribution policy, up to the private equilibrium margin m_1, is the cost ABC in Figure 1. The cost is then the integral from m_o to m_1 of the extra mileage ($2m_1$-2r) times cost per mile c times number of standard trips per person τ times population at that distance $D_m 2\pi r\,dr$. Assuming 250 work days per year, the yearly cost is obtained by multiplying by 250 to obtain

$$1000\,\pi c\tau\, D_m \int_{m_o}^{m_1} (m_1 - r)\, r\, dr$$

which equals

$$1000\,\pi c\tau\ D_m\ m_1\ (m_1^2 - m_o^2)/2 - (m_1^3 - m_o^3)/3 \quad .$$

Suppose the edge m_o to which the city is constrained is 25 miles from the central business district. Under the assumption that additional growth would take place at suburban density of 1000 persons per square mile in a circular expansion, the area required for a growth of population of 600,000 would be such that the added area times the density would equal the added population, the condition is $(\pi m_1^2 - \pi m_o^2)\ 1000 = 600{,}000$, or $m_1 = (600/\pi + m_o^2)^{1/2}$. Given a value of m_o of 25, the solution for m_1 is 28 miles.

All the needed values have now been given. Inserting the values into the foregoing centered expression indicates as the area ABC a yearly loss of $43 million, due to the movements along private costs curves if growth of the city is restricted by 10 percent.

Density

While city scale effects require attention to slope of the privately perceived cost curve determining the area ABC as just considered, most of the analysis of effects of population redistribution involves external or unperceived costs making a difference between height of private and social curves and determining the area ABDE. The external or unperceived costs will be of concern in the remainder of the paper.

As was pointed out in the framework section, one reason for difference between private and social costs is effect of city growth on density. Starting from very small towns and contemplating larger ones, average density of a town must rise as the increasing distance to work makes residential land near work places more valuable, inducing the construction of denser housing including multi-family structures. Whether average density continue indefinitely to increase is moot. For an existing large city, the new growth on the city's edge will be at a suburban density which is exceeded in increasing magnitude by the higher densities prevailing nearer the center. Only if the further increase in density nearer the center induced by city growth overcomes the decrease due to the lowering effect on average density of the growth at the edge, will average density of the city increase. Buildings already in place which would have to be torn down to increase density will impede the density response. The possibility of development of new sub-centers within the city, instead of continuing to increase the densities in proximity to existing centers and sub-centers, could further retard average density response. For these reasons, one at the least expect an eventual slowing in the increase in average density as a city grows, and conceivably there could be a fall.

For a city with a single center, the average density is determined by

a density gradient beginning at the suburban density at the edge of the city and rising continuously up to maximum density at the center. If the city is not too large, density may rise by a constant exponential amount over the entire range:

$$D_r = D_m e^{kr}$$

where D_m is density at the edge, m is the size of the city measured as the distance between the edge and the center, r is the distance from the edge going toward the center of the city and D_r is density measured as persons per square mile at the distance r.[1] If the city is circular, the population in a doughnut shaped ring is $D_r\ 2\pi r\ dr$ or substituting out the expression for D_r just given, $D_m\ e^{kr}\ 2\pi r\ dr$. Summing over all the rings between the margin and the center, the total population N of the city is $N = \int_0^m 2\pi D_m e^{kr} r dr$. Carrying out the integration gives

$$N = 2\pi D_m\ (e^{km} - 1 - k_m)/k^2$$

Dividing population N by the city area A gives the average density of the city $\bar{D}$, or $\bar{D} = N/A$. To find out how average density is affected by city growth, differentiate $\bar{D}$ with respect to N and multiply by $N/\bar{D}$ to obtain the percentage change in average density resulting from a one percent increase in population, $(d\bar{D}/dN)(N/D) = 1 - (dA/dN)\ (N/A)$ which is the percentage increase one in population minus the percentage change in city area. The problem thus reduces to finding the effect of population growth on city area. Again assuming the city is circular, its area A is $2\pi m^2$. Differentiating the area with respect to m and then multiplying by N/A reveals that the percentage increase in area resulting from a one percent increase in population $(dA/dN)(N/A)$ is $2\ (dm/dN)(N/m)$. Substituting this result into the equation for $(d\bar{D}/dN)(N/D)$, the percentage change in average density is

$$(d\bar{D}/dN)(N/D) = 1-2\ (dm/dN)(N/m)$$

To find the percentage increase in the distance m between the margin and the center as a result of the one per cent increase in population, take the differential of the equation for population $N = 2\pi D_m(e^{km}-1-km/k^2$ letting N and m vary. The result is $dN = (2\pi D_m/k)(e^{km}-1)dm$. Solving for dm/dN and multiplying by N/m gives $(dm/dN)(N/m) = 1/k_m - 1/(e^{km}-1)$ which inserted into the foregoing centered expression reveals the percentage change in average density resulting from a one per cent change in population to be

$$(d\bar{D}/dN)(N/D) = 1 - 2\{[1/km] - [1/(e^{km}-1)]\}.$$

Because of the possibility that the assumed shape of the density function is not suitable for predicting effects of growth of a very large city, a numerical example will be given for a city of one million persons.

---

[1] For empirical evidence on density gradients see Richard Muth, Cities and Housing (Chicago: University of Chicago Press, 1969).

In the formula for population $N = 2\pi D_m(e^{km} - 1 - km)/k^2$, a population of about one million will be obtained by assuming a suburban density D_m of 1000 persons per square mile, increase in density per mile k going toward the center of .2 and distance m from margin to center of about 10 miles. Applying the value of km of 2 in the expression just given for $(d\bar{D}/dN)(N/D)$ indicates that the elasticity or percentage change in density resulting from a one percent change in population is .3.

The next task is to find the effect of change in density on costs. While density is often found to be significant in studies of local government expenditures, systematic investigation of effects of density on costs has apparently not been undertaken. For purposes of a numerical example, a reasonable value of the elasticity of costs with respect to density is -.1. This is not an exact estimate but is a judgment suggested as being realistic, based on studies of waste collection costs and of education expenditures undertaken at the University of Chicago and on a review of other studies reported by Werner Hirsch.[1]

Combining the foregoing results, the elasticity of percentage change in costs resulting from a one percent increase in population, due to effects on density, is the elasticity of costs with respect to density times the elasticity of density with respect to population, that is -.1 times .3, or -.03. The costs affected include all production involving significant daily travel in the city such as commuting already considered, much business pickup and delivery and a variety of local services provided by governments. For purposes of the example, suppose that one third of the income produced in the city is subject to cost reduction. Then the elasticity of total costs, including both the commodities affected and those not affected, is one third of -.03 or -.01. If the total income produced in the city is $2.5 billion, the effect on production costs of growth of the city population by 10 percent or one-tenth, is one tenth times -.01 times $2.5 billion or a cost savings of $2.5 million. In comparing with the example of the previous section and examples to be given later, it should be remembered that the result here is for a city of 1 million while the other examples are for a city of 6 million.

Congestion

As one of the undesirable external costs associated with growth of large cities, which would affect area ABDE in Figure 1, the increase in traffic congestion costs resulting from an increase in city size was considered in a previous paper.[2] An illustration was given of a congestion cost of $.10 per mile driven in a large city, amounting to $.50 per day per worker. The illustration assumed a linear speed-flow relationship and did not consider road capacity decisions. To extend that analysis, the effects of a nonlinear

---

[1] Hirsch, Werner, The Economics of State and Local Government (New York: McGraw-Hill, 1970).

[2] G. S. Tolley, "Population Distribution Policy," presented at Public Policy Education Conference, Custer, South Dakota, August, 1971.

speed-flow relationship and non-optimal road capacity decisions may be considered.

The cost of an auto trip of length D for the person deciding to make the trip, is the car operating cost plus the time cost of the trip or

$$p = D(c+w/s)$$

where p is cost per trip, c is car operating cost per mile, w is the wage or other amount at which the time cost per hour of travel is valued, and s is speed of travel.

The number of auto trips made at times of day when there is traffic congestion depends on the cost of making the trips. People can vary the number of auto trips made under congested conditions by rescheduling travel to uncongested times, travelling by train or bus and reducing total trips taken. At higher costs per trip, people will be increasingly willing to bear the alternative costs encountered in reducing auto trips at congested times. The demand schedule for auto trips at congested times can be expressed as

$$p = p(x),$$

showing cost per trip corresponding to a given x, whereas x is number of trips per person.

Assuming the normal situation of collecting no congestion toll, a first condition specifies that the number of trips taken per person will be the number of trips on the demand schedule indicated by the cost for a person deciding to make trips or

(1) $D(c + w/s) + t = p(x)$

where t, the congestion toll per trip, is assumed to be zero.

The speed of travel s which affects costs depends on the total number of trips taken by all persons. Speed also depends on characteristics of the road network such as the number of roads, their width and arrangements provided for avoiding delays due to intersecting traffic. The relation determining speed gives a second condition and can be written

(2) $s = s(v,k)$

where

(3) $v = nx.$

The symbol n refers to the total number of persons in the city, so that v is total number of trips taken in the city under congested conditions. The total amount of resources devoted to providing the road network is denoted

as k. Conditions (1), (2) and (3) jointly determine x (number of trips taken per person), s (speed of travel) and v (total number of trips), population of the city n, the road network characteristics as determined by k and the cost parameters D and w.

Knowledge of the relation determining speed enables estimating the cost imposed by taking a trip which a person deciding to take the trip does not pay for. To find the cost not paid for, which is due to the slowing down of all other cares, note that the total cost of travel is traffic volume times cost per trip or vD(c+w/s). The cost of an extra trip is the derivative of total cost with respect to traffic volume d vD(c + w/s /dv or D(c + w/s) - $(vDw/s^2)(\partial s/\partial v)$. The first part of this expression is the cost to the driver already noted. The second part of the expression is the cost not paid for. It depends on the partial derivative $\partial s/\partial v$, which is the effect of a unit increase in traffic volume on speed resulting from a one per cent increase in traffic volume. Since e_v equals $-(\partial s/\partial v)(v/s)$, and since the cost may be expressed on a per mile basis by dividing by length of trip D, the cost not paid for is $(w/s)e_v$ per mile of travel. In words this is the time cost of travelling a mile multiplied by the volume elasticity.

The general nature of the speed relation is shown in Figure 2. Each curve corresponds to a given amount of expenditures on the road network and shows how speed s declines at higher values of traffic volume v. The range of no congestion accounts for the flat part of each curve. Speed is reduced as volume rises above the point where congestion is encountered. If the effect of extra cars on speed, i.e. marginal effect of cars, increases gradually from zero to a maximum effect, as the volume rises beyond the point of no congestion, each curve will display an increasingly negative slope as depicted in the figure. Algebraically the assumption is that the absolute value of $\partial s/\partial v$ increases with rising volume. For a given road network, there is a maximum traffic volume at which speed is effectively reduced to zero. This is the point at which each curve reaches the horizontal axis in Figure 1.

As equation (2) states, in addition to being affected by traffic volume, speed is affected by road expenditures k. Figure 1 assumes that increased road expenditures raise speed in the absence of congestion, maximum un-congested traffic volume and the volume at which speed is finally reduced to zero.

In the analysis of the earlier paper, the speed relationship was approximated linearly by passing a line through an estimated noncongested speed and zero volume (the point L in the figure) and the observed speed and volume (the point N). The estimated slope of the speed relation was then the slope of the line LN. If the assumptions of Figure 2 are correct, this approximation underestimates the negativity of the slope of the speed relation, leading to an underestimate of the volume elasticity and hence to an underestimate of the cost a car imposes on others. The example in the previous paper may be regarded as providing a method for making a lower limit estimate

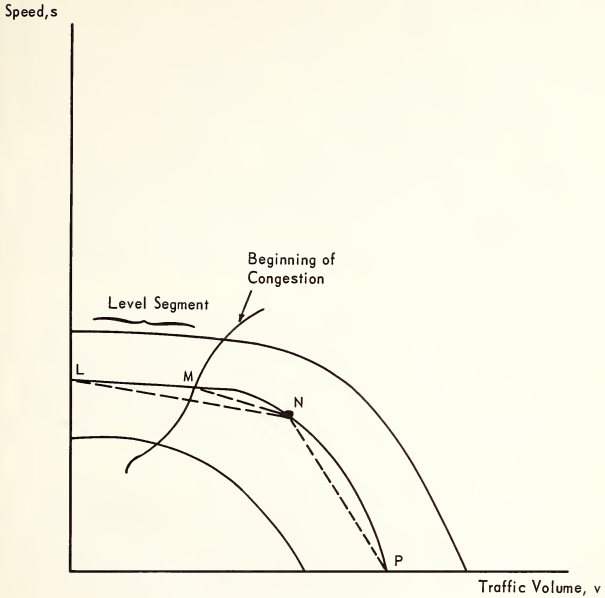


FIGURE 2. Speed-Volume Relationship

of congestion costs. A better approximation, though still an overestimate, would be obtained by considering a linear approximation involving the excess of traffic volume over that at which congestion begins. The assumption would be that speed is reduced linearly beginning at the point M at which congestion is first encountered. The estimated slope of the speed relation is then the slope of the line MN. The speed at point M is the same as at point L, but the volume at M is the maximum noncongested volume instead of being zero. Suppose the estimate were that the maximum noncongested traffic volume was one half the observed volume. The absolute value of the slope of MN would be twic that of LN. The estimated volume elasticity would be doubled, and the estimated cost not paid for would also be doubled. In the earlier example, the noncongested speed was 35 miles per hour, and the observed speed s was 20 miles per hour. The slope of LN as the estimate of $\partial s/\partial v$ was $(35-20)/v$, giving an estimate of e_v of $[(35-20)/v]\ (v/20)$ or .75. Using the slope of the line MN, the estimate of e_v is $[(35-20)/(v-.5v)]\ (v/20)$ or 1.5. Given the assumption of a value of travel time per car of $3.00 per hour, the earlier estimate of the cost not paid for was about $.10 per mile. Under the new assumption this would be raised to about $.20 per mile. This refinement does not change the conclusions of the earlier paper.

While the slopes of LN and MN give underestimates of the volume elasticity, considering the point P at which speed is reduced to zero can give an overestimate. The slope of the line NP overestimates the slope of the speed relation. Suppose speed would be reduced to zero at a traffic volume 50 percent greater than that observed. The slope of NP is equal to the observed speed divided by the excess of traffic volume at zero speed over observed traffic volume, or slope equals $[s/(1.5v - v)]$. The estimated volume elasticity e_v is then $[s/(1.5v-v)]\ (v/s)$ or 2. The estimate of the cost not paid for becomes $.30 per mile.

This discussion has shown how the estimate of the cost not paid for can be bracketed by estimates of the two key points N and P on the speed-volume curve. For purposes of illustration, assume that the congestion cost is midwa between the $.20 and $.30 estimate, or $.25 per mile. If there is an average per worker of 5 miles each work day of driving under congested conditions and if there are 250 work days per year, a worker incurs an extra travel cost not paid for each year of $312.50. If restricting the growth of a city of 6 million persons by 10 percent excludes 600,000 people, one-quarter of whom are workers, the savings is $312.50 times 150,000 or about $47 million per year.

This order of magnitude estimate has been developed without considering road building expenditures k. If a person entering a city does not change road building expenditures, then the external effect in slowing others down corresponds exactly to the optimal congestion tax as assumed in the estimate. However, ordinarily as a city grows more road building will be undertaken. The extent of this road building will determine how speed of traffic is affected by city growth. In terms of Figure 2, city growth involves a shift to a new speed volume curve and is not confined to movements along one curve.

Another simplification of the estimate just developed is that it takes as given the amount of automobile travel per person under congested conditions. Speed s affects price per trip to people travelling by automobile, so that changes in speed may lead to a price response as determined by the demand function p(x). The notion widely believed that traffic congestion per person increases with city size, if true, indicates a tendency for road resources k not to be increased sufficiently to prevent declines in travel speeds as city size increases. A more general framework will now be developed allowing for changes in road building expenditures and effects of speed on automobile travel.

Conditions (1) through (3) already presented are three of four relations needed for a more general framework. Condition (1) which is $p(x)=D(c+w/s)+t$ specifies the number of trips as influenced by the price of trips. Condition (2) is the relation $s=s(v,k)$ showing how speed is affected by traffic volume v and road expenditures k. Condition (3) or $v=nx$ defines traffic volume as the product of population times trips per person. The fourth condition specifies the amount of road building expenditures and has not yet been considered.

It is reasonable to suppose that road expenditures are influenced by the marginal benefits from them. This can be done without assuming that an exact optimization is achieved. The benefits from spending an extra dollar on roads consist of the savings on existing trips due to faster travel plus the benefits, if any, from the extra trips induced by the spending. The extra trips induced by the expenditures slows traffic down compared to speed if traffic volume were not affected. These extra trips will take away from benefits unless the value of the extra trips exceeds their cost sufficiently to make up for the slowing down effect.

To consider more explicitly how a city of a given size decides on road expenditures, differentiate total travel costs $nxD(c+w/s)$ with respect to k and change sign to express as a benefit, i.e. saving in travel costs, $(nxDw/s^2)(ds/dk)$. The effect on speed ds/dk when extra road building expenditures are incurred depends on the speed relation $s=s(v,k)$. Differentiating this relation with respect to k gives $ds/dk=ns_v(dx/dk)+s_k$ where $v=nx$ or $dv/dk=n(dx/dk)$ has been used in carrying out the differentiation. The term s_k is the effect on speed of extra road expenditures in the absence of any effect on number of trips, whereas $ns_v(dx/dk)$ is effect on speed of extra trips induced by the expenditures. Using the result for effect on speed in the expression for saving in travel costs gives as the saving in travel costs the decrease in costs on existing trips, assuming total trips are not affected, minus the increase in costs due to the slowing down effect of extra trips, or $(nxDw/s^2)s_k+(n^2xDw/s^2)s_v(dx/dk)$. The second term is negative indicating an increase in travel costs since s_v is negative. Using the definitions of elasticities of speed with respect to road expenditures and traffic volume given above, the saving in travel costs is $(nx/k)(Dw/s)e_k - (nDw/s)e_v(dx/dk)$. Against the second term indicating

travel costs imposed by speed effect of extra trips must be offset the benefit if any of the extra trips. The benefit of the extra trips is the difference between the value of the trips and the travel costs incurred to take them or n p(x)-D(c+w/s) (dx/dk). In turn, from condition (1), the value of the trips is p(x)=D(c+w/s)+t. Making this substitution for p(x), the benefit of the extra trips is nt(dx/dk). Combining the saving in travel costs with the benefits of the extra trips gives as the marginal benefit of road expenditure $(nx/k)(Dw/s)e_k+n[t-(Dw/s)e_v]$ (dx/dk). This is the same result as can be obtained more simply by differentiating with respect to k equation (21) given below which is the total benefit of the road.

To evaluate the marginal benefit expression just given requires knowing dx/dk, the change in trips induced by an increase in road building expenditures. A change in road building expenditures affects the speed of traffic, reducing the costs of travelling to motorists and inducing them to take more trips. The change in trips induced by road building expenditures thus depends on the speed relation, the effect of speed on trip cost and the trip response to cost as determined by the demand function. The change dx/dk in trips resulting from a one dollar increase in road expenditures times the change dp(x)/ds in trip cost per unit increase in speed times the change dx/dp(x) in trips taken per unit increase in trip cost. That is, dx/dk=(ds/dk) [dp(x)/ds] [dx/dp(x)]. The first of these derivatives ds/dk, itself depends on the change in number of trips since the volume of traffic affects speed as determined by the speed relation. As already noted, differentiating the speed relation s=s(v,k) with respect to k gives $ds/dk=s_v n(dx/dk)+s_k$. The change in trip cost to motorists dp(x)/ds from (1) is $d[D(c+w/s)+t]/ds$ which equals $-Dw/s^2$, if the realistic assumption is made that there is no change in congestion tax. The change dx/dp(x) is the effect of a unit change in price of trips as determined by the slope of the demand schedule for trips. Multiplying the three derivatives together to obtain an expression for dx/dk gives an equation for dx/dk where dx/dk also appears on the right hand side. Solving for dx/dk and simplifying gives

$$dx/dk = -(x/k)\ \ e_k/(-e_v+1/\rho\beta)$$

where e_k and e_v are the elasticities of speed with respect to road expenditures and traffic volume previously defined, ρ is (w/s)/(c+w/s) or time cost as a fraction of total travelling cost and β is the elasticity of demand for trips with respect to their price. The change dx/dk just derived is governed by conditions (1) through (3) determining traffic flows. A road planner needing to know the effect of extra road expenditures on trips could, if he knew the relations, differentiate them with respect to k obtaining three equations in the derivatives dx/dk, dv/dk and ds/dk. The solution of the three equations for dx/dk is the expression just given.

Substituting the expression for dx/dk into the expression for marginal benefit of road expenditures gives $(nx/k)(Dw/s)e_k-(nx/k)\ [t-(Dw/s)e_v]$

$[e_k/(-e_v+1/\rho\beta)]$. A road building planner can be visualized to compare this benefit from spending a dollar with the one dollar cost. With perfect optimization, the marginal benefit and cost would be equated by setting their difference equal to zero. There are several reasons for supposing that roads are not always built to this point, either falling short of the so-called optimal capacity or going beyond it. Road expenditures are decided by a public process, and the benefits are at best difficult to evaluate precisely. Tax and borrowing decisions are required at the local and state levels, imparting a tendency for rich communities to overbuild roads and poor communities to underbuild them. Expenditures are influenced by attitudes reflected in the public decision-making process. Growth minded communities may tend to build beyond the point suggested by the criterion. Communities not desiring growth may deliberately underbuild. Subsidies for expressways can give a community an incentive to build beyond the point suggested by the criterion.

Let m be the difference between the marginal benefits and marginal cost of extra road building expenditures. Then the relation being sought specifying road building expenditures. Then the relation being sought specifying road building expenditures k is

(4) $(nx/k)(Dw/s)e_k - (nx/k)\left[t-(Dw/s)e_v\right]\left[e_k/(-e_v+1/\rho\beta)\right] - 1 = m,$

that is, the benefit of an extra dollar of expenditure as derived up to this point, given by the first two terms, minus the one dollar cost, equals the difference m between marginal benefit and cost. Under certain assumptions, condition (4) reduces to the simpler condition that the savings in expense on existing trips should equal the extra road building cost or $(nx/k)\ (Dw/s)e_k-1=0$. One assumption required is perfect optimization or m=0. The simpler condition would still only be obtained if there were no effect on trips, i.e. $\beta = 0$, or if the optimal congestion toll were charged, i.e. $t=(Dw/s)e_v$. In reality neither of these latter conditions is fulfilled.

Conditions (1) through (4) determine trips per person, total number of trips, speed and road building expenditures for a city of given population. We are now in a position to consider how these magnitudes are affected by a change in population. Taking the differentials of the four conditions, dividing through by absolute values and making divisions and multiplications as necessary to obtain elasticity expressions gives

(1') $\qquad x/\beta = -\rho\,\dot{s}$

where β is price elasticity of demand for trips $p/xp'(x)$ and ρ is time cost of a trip as a percent of total cost of a trip or $(Dw/s)/\left[D(c+w/s)+t\right]$;

(2') $\qquad \dot{s} = e_v\dot{v} + e_k\dot{k}$

where e_v and e_s are the elasticities of speed with respect to volume and

expenditures previously defined;

(3') $\dot{v} = \dot{n} + \dot{x}$

and

(4') $(1+e_{kv}+\varepsilon)\dot{v} + (e_{kk}+\varepsilon e_{vk})\dot{k} - (2+\varepsilon)\dot{s} = 0$

where e_{kv} and e_{kk} are the elasticities of the marginal effect of road expenditures on speed with respect to traffic volume and road expenditures

or $e_{kv} = vs_{kv}/s_k$ and $e_{kk} = ks_{kk}/s_k$, e_{vk} is the elasticity of the marginal effect of volume on speed with respect to road expenditures or $e_{vk} = ks_{vk}/s_v$ and ε is a negative number whose absolute value is less than one and specifically is $\varepsilon = \rho\beta e_v/(1-\rho\beta e_v)^2$.

In taking these differentials it has been assumed that t is zero which is realistic inasmuch as congestion tolls are essentially nonexistent. It has also been assumed that $\dot{m}$ is zero. This assumption will definitely be fulfilled if road expenditures are optimal (m=0). Equation (4') is also consistent with departures from optimality. The assumption that $\dot{m}$ is zero implies there is no systematic change one way or the other in difference between marginal gain and marginal cost of road expenditure as city size increases. Economies or diseconomies of scale, in connection with average cost financing, might lead to $\dot{m}$ not being zero. Economies and diseconomies will be considered later in this section. The assumptions that t and $\dot{m}$ are zero could be dropped by adding two more equations explaining t and $\dot{m}$ in terms of $\dot{n}$.

Equations (1') through (4') determine the percentage changes in trips per person, speed, traffic volume and road expenditures ($\dot{x}$, $\dot{s}$, $\dot{v}$ and $\dot{k}$) given the percentage change in population $\dot{n}$. Equations (1') through (4') can be solved by any method applicable to linear equations, such as determinants or use of equations one by one to eliminate variables. The solution for the percentage change in speed is

(5) $\dot{s} = \dot{n}\left[e_v(e_{kk}+\varepsilon e_{vk}) - e_k(1+e_{kv}+\varepsilon)\right] / \left\{-\rho\beta + \left[e_{kk}+\varepsilon e_{vk} - e_k(2+\varepsilon)\right]\right\}$.

The other endogenous percentage changes can be similarly obtained if needed.

The special case of no change in trips per person is obtained if the elasticity of demand for trips β is zero, in which case ε is also zero. Inserting these conditions into (5), the percentage change in speed becomes

(6) $\dot{s} = \dot{n}\left[e_v e_{kk} - e_k(1+e_{kv})\right] / (e_{kk} - 2e_k)$, if $\beta = 0$.

Another special case occurs if no changes in road building expenditures are allowed as city size changes. Then $\dot{k}$ is zero, and equation (4') is dropped. The solution for the percentage change in speed is

(7) $$\dot{s} = \dot{n}\left[ev/(1-p\beta e_v)\right], \text{ if } \dot{k} = 0.$$

A final special case occurs when there is no change in trips per person and road capacity does not change. Letting $\beta=0$ in the foregoing equation gives percentage change in speed

(8) $$\dot{s} = \dot{n}e_v \text{ , if } \beta=0 \text{ and } k=0.$$

To further examine changes in speed, the following functional form for the speed relation may be considered:

(9) $$(s/\bar{s}) = 1 - a\left[(v/\bar{v}) - 1\right]^{\gamma}$$

where $\bar{s}$ is maximum uncongested speed attained, when traffic is light enough so that cars do not slow each other down, and $\bar{v}$ is the maximum uncongested traffic volume, above which the cars slow each other down. Both $\bar{s}$ and $\bar{v}$ may be increased by road building expenditures, i.e.

(10) $$\bar{s} = \bar{s}(k)$$

and

(11) $$\bar{v} = \bar{v}(k)$$

where $\bar{s}'(k)$ and $\bar{v}'(k)$ are positive.

Taking partial of (9) with respect to v and using the substitution from (9) that $-a\left[(v/\bar{v})-1\right]^{\gamma} = (s/\bar{s})-1$ gives $\partial s/\partial v = \gamma(s-\bar{s})/(v-\bar{v})$. Using this result in the definition of e_v of $-v(\partial s/\partial v)/s$ gives

(12) $$e_v = \gamma\left[\bar{s}/s)-1\right] / \left[1-(\bar{v}/v)\right].$$

Substituting (10) and (11) into (9), differentiating with respect to k, and eliminating the exponential through substituting expressions for $(s/\bar{s})$ and $(s/\bar{s})-1$ obtained from (9) gives $\partial s/\partial k=(s\bar{s}'/\bar{s}) + (v\bar{v}'/\bar{v})(\bar{s}-s)/(v-\bar{v})$. Using this result in the definition of e_k of $K(\partial s/\partial k)/s$ gives

(13) $$e_k = \eta + \pi\gamma\left[\left[\bar{s}/s)-1\right] / \left[(1-(\bar{v}/v)\right]\right.$$

where $\eta=k\bar{s}'/\bar{s}$ is the elasticity of uncongested speed with respect to road expenditures and $\pi=k\bar{v}'/\bar{v}$ is the elasticity of maximum uncongested traffic volume with respect to road expenditures. The values of η and π are determined by (10) and (11), as $\bar{s}'$ is the first derivative of (10) and $\bar{v}'$ is the first derivative of (11).

To find e_{vk} differentiate $\partial s/\partial v$ with respect to k, use this result in the definition $e_{vk}=ks_{vk}/s_v$, and multiply and divide as needed to obtain elasticity expressions. The result is

$$(14) \qquad e_{vk} = \gamma/\left[1-(\bar{v}/v)\right] + \eta/\left[1-(s/\bar{s})\right] + \pi/\left[(v/\bar{v})-1\right]$$

The analogous procedure may be followed using the definition $e_{kv}=ks_{kv}/s_k$ to obtain

$$(15) \qquad e_{kv} = (e_v/e_k)\left\{\left\{\eta+\pi\left\{1-e_v/\left[(\bar{s}/s)-1\right]-1/\left[1-(\bar{v}/v)\right]\right\}\right\}\right\}.$$

In obtaining $e_{kk}=ks_{kk}/s_k$ the second derivatives $\bar{s}''(k)$ and $\bar{v}''(k)$ are encountered. If to obtain an expression for e_{kk} one assumes constant elasticity functions for (10) and (11), which has the advantage of allowing for diminishing returns to road building expenditures, η and π will be invariant in the differentiation giving

$$(16) \quad e_{kk}=(\eta/e_k-1) + (\pi e_v/e_k)\left\{\pi/\left[1-(s/\bar{s})\right]-e_k/\left[(\bar{s}/s)-1\right]+\pi/\left[(v/\bar{v})-1\right]\right\}.$$

Note that the parameters must be consistent with observed speed and traffic volume. Under the assumptions of the earlier example, that observed speed s is 20 while uncongested speed $\bar{s}$ is 35, and that uncongested traffic volume $\bar{v}$ is one half the observed volume so that $v/\bar{v}$ is 1/.5 or 2, (9) implies $(20/35)=1-a(2-1)^{\gamma}$. This gives a value of a of 15/35. Still following the earlier example, assuming speed would be reduced to zero at a traffic volume 1.5 times the observed volume, then at zero speed the ratio of actual volume to noncongested volume would be 1.5/.5 or 3. Insertion of the zero speed conditions into (9) gives $0=1-(15/35)(3-1)^{\gamma}$ or $\gamma=(\ln 35-\ln 15)/\ln 2$. The value of γ is 1.22.

This value of γ together with the conditions for observed traffic $\bar{s}/s=35/20$ and $\bar{v}/v=.5$ inserted into (12) gives $e_v=1.83$. Consider the increase in road expenditures that would be necessary to maintain the same speed if traffic volume were increased. The assumption that a 10 percent increase in volume would require a 10 percent increase in road expenditures to maintain the same speed implies that e_k equals e_v. Note that, using (12), (13) can be written $e_k=\eta+\pi e_v$. With $e_k=e_v=1.83$, the assumed values of η and π must then be consistent with $1.83=\eta+1.83\pi$. As road expenditures are increased, there is some optimum division of expenditures as between increasing uncongested speed and traffic volume possible before congestion is encountered. This division determines the relative values of η and π. Intuitively, it seems likely that most effort in the normal situation would be devoted to increasing the capacity of roads and streets to handle traffic at existing speeds, rather than attempting to increase maximum speed as city size increases. In obtaining a numerical result here, it will be assumed that all additional road expenditures are devoted to increasing traffic volume with no increase in maximum speed, i.e. $\eta=0$, $\pi=1$. Using (14), (15) and (16), the following values are then obtained: $e_{vk}=3.44$, $e_{kv}=.344$ and $e_{kk}=-1.44$.

The values in the preceding paragraph can be used to estimate how speed will change as city population increases. From (8), if the elasticity of demand β is 0 and if road expenditures are not allowed to change, the percentage change in speed resulting from a one percent increase in population $\dot{s}/\dot{n}$ is -1.83. From (7), if β is 1 and the ratio ρ of time cost to total travel cost is .5, still retaining the assumption that road expenditures do not change, $\dot{s}/\dot{n}$ is -.96. From (6), if β is zero but road expenditures are allowed to change, $\dot{s}/\dot{n}$ is -.36. From (5), if β is 1 and road expenditures are allowed to change, $\dot{s}/\dot{n}$ is -.11.

To use these results to examine the effects of population redistribution on national income, let y refer to all income earned by a person in the city in question other than the net benefits of using the city's congested roads. The total income earned by a person in the city if y plus the net benefits of the person b from his use of the congested roads, or y + b. Let y' refer to the income the person would earn if he were outside the city in question, where y' is all income from being elsewhere including benefits from using any roads in other cities. Labor will have incentives to move until the benefits from an extra person moving to the city in question are zero. The income gained from moving to the city is y+b, and the income given up is that available elsewhere or y'. Thus incentives foster the condition

(17) $$b + y - y' = 0.$$

This incentive will result in maximum national income, i.e. the national income gains from the movement of an extra laborer being zero, only if the change in benefits from congested road use dB/dn plus other income y earned in the city equals total income y' the person could earn elsewhere, assuming y and y' are income produced with no externalities. For purposes of the present section concentrating on external effects connected with road use, it is appropriate to assume that what is produced in this city other than benefits of using congested roads equals y and what would be produced elsewhere equals y'. Divergencies from these conditions concerning y and y' are the concern in the analysis elsewhere in this paper of externalities other than those connected with road use. Under the assumed conditions, the effect on national income when a person moves to the city is

(18) $$dB/dn + y - y'.$$

Substituting the difference between y and y' expected in view of market incentive conditions given in (17), the effect on national income when a person moves to the city is

(19) $$dB/dn - b,$$

that is, the difference between the marginal effort of a person on road benefits and the benefit b a person receives from the road when he moves to the city.

Confining attention to that part of the road system which is congested, the benefit to a person from using the roads when he moves to a city is

$$(20) \qquad b = \int_0^x p(X)dX - xD\,(c+w/s) - k_n.$$

The first term on the right hand side is the area under the demand curve of a persons to travel on the roads from zero trips up to the actual number of trips taken. It is the sum of the amounts that he would be willing to pay to use the road system and is greater than his travel expenditures due to willingness to pay higher prices per trip at smaller numbers of trips. The second term on the right hand side is the travel cost borne by a person using the roads. It is the cost per trip $D(c+w/s)$ times x, the number of trips per person. The third term k_n is the person's contribution to road expenditures. It is the amount his taxes are higher because of provision of the roads.

The total benefits from using the roads is

$$(21) \qquad B - n\int_0^x p(X)dX - nxD\,(c + w/s) - k.$$

The first two terms on the right side of (21) multiply by population n the first two terms pertaining to travel benefits and costs on the right side of the expression for individual benefits (20), and the third term k is the total cost of providing the roads. To find the effect on the road benefits of an extra person entering the city, differentiate (21) with respect to n:

$$(22) \qquad dB/dn = \int_0^x p(X)dX - xD(c+w/s) - dk/dn$$

$$+n\left[p(x) - D(c+w/s)\right](dx/dn) + (nxDw/s^2)(ds/dn).$$

In this example it will be assumed that a person's contribution to road expenditures is equal to the added cost in road building expenditures which his presence in the city causes. If this is not so, there is an institutional externality whereby a difference between benefits from public expenditures and a person's taxes leads to migration incentives departing from incentives for labor allocation maximizing national income. Institutional externalities for roads will be considered at the close of this section, and institutional externalities will be discussed as a more general phenomenon in the closing section. Assuming for the moment no institutional externality connected with payment for road expenditures implies $dk/dn=k_n$. With this assumption, the first three terms of (20) and (22) are identical. The bracket in the fourth term of (22) is the difference between the price of the trip and its cost to the driver, or zero according to the condition of road use given in equation (1). Under these assumptions, subtracting (20) from (22) to find effect on national income when a person moves to the city as indicated by (19) gives

$$(23) \qquad dB/dn - b = (xDw/s)\dot{s}/\dot{n},$$

where the percentage change notations $\dot{s}$ and $\dot{n}$ have been substituted for

ds/s and dn/n.

Numerical examples may be obtained by multiplying the results for $\dot{s}/\dot{n}$, obtained from (5) to (8) by xDw/s. The magnitude xD is miles per person of congested driving. Following the earlier example, 5 miles per day times 250 work days is 1250 miles of congested driving per year. If there are four persons per worker, xD is 1250/4 or 312.5 miles of congested driving per person per year. The magnitude w/s is the time cost of driving a mile estimated as a wage divided by speed. With the assumptions of the earlier example, w/s is $3.00 per hour divided by 20 miles per hour or $.15 per mile. Thus xDw/s, which is seen to be the yearly time cost of congested driving per person in the city, is $46.875. Multiplying xDw/s by the results for $\dot{s}/\dot{n}$ obtained above under alternative assumptions (5) to (8) gives the effect of growth of the city by one person. Further multiplication by 600,000 gives the effect of growth of a city of 6 million persons by 10 percent.

The estimated effect is $51 million if $\beta=0$ and k=0. This corresponds to the assumptions of the original example and is very close to the $47 million estimate given initially. The only reason for difference is that the initial estimate used as the estimate of speed elasticity e_v an interpolated value, whereas the present result uses a point on an assumed functional form. Applying other assumptions, the effect is $27 million if $\beta=1$ and k=0, $10 million if $\beta=0$ and road expenditures are allowed to change, and $3 million if $\beta=1$ and road expenditures are allowed to change.

The examples suggest that losses due to congestion, as city size is increased, are reduced by nonzero elasticity of demand and by changes in road expenditures. Further refinements are possible and might suggest reasons for obtaining higher instead of lower losses as compared to the original example. All the analysis in this section has assumed a constant speed under congested conditions and has assumed that city growth does not affect number of miles of congested driving. Suppose congestion is due to slow-down within a distance of 2.5 miles from work (encountered once in the morning and once in the evening for a total of 5 miles per day). Figure 3a depicts congestion costs if the congested speed is constant regardless of distance from work, with congestion cost per mile plotted on the y-axis and distance from work plotted on the x-axis. The area R is the total congestion cost per trip to or from work, and the area Q is the increase in costs due to growth of the city as estimated above.

Figure 3b depicts a more realistic situation where congestion is encountered at the same distance, but speed is slower the closer one is to the place of work. For simplicity linear effects are depicted in Figure 3b but the argument is the same if the lines are curved. The line rr showing congestion cost per mile would be much more steeply sloped if it were not for road expenditures undertaken to hasten traffic flow near the place of work. Extra expenditures are to be expected due to increasing traffic volume near work place, as persons living nearer and nearer to work enter the road system. On the other hand, one would not expect it to be optimal to completely flatten the travel speed gradient. These ideas are corro-

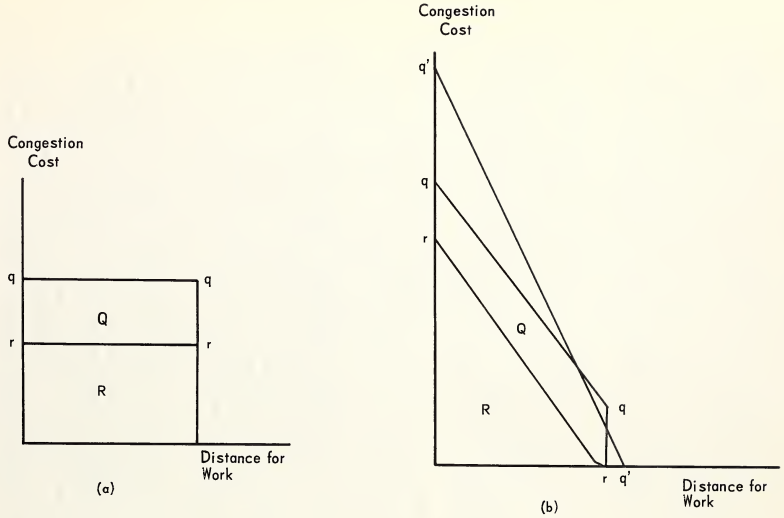


FIGURE 3. Effects of Added Population on Congestion Costs

borated by common experience. While there are more elaborate roads near the center of town, traffic is also slower there.

As long as the area R in Figure 3a is the same as the area R in Figure 3b, no bias in estimating total congestion costs results from the unrealistic constant speed assumption. However, eliminating the bias in estimate of congestion costs imposed by additional population could be more difficult than eliminating bias in estimate of total congestion cost. If congestion costs imposed by additional population were a constant amount per mile of road already congested, as indicated by line qq in Figure 3b being the same distance above rr at all points as in Figure 3a, the method of estimating costs of additional population developed above would be adequate. The more likely situation is that the congestion costs with added population are depicted by q' q'. Congestion is encountered at a greater distance, and the increase in congestion is greater the closer one comes to the place of work. The basic reason for the position of q'q' is the point already made that extra road expenditures are undertaken to partly, but not completely, reduce congestion. One would expect the largest increases in road expenditures in the places where the increases in congestion are greatest, but one would not expect the greater congestion increases to be completely eliminated. The more the increase in congestion, the greater will be the expenditures to reduce congestion but the greater also will be the increase in congestion remaining. Given the nature of the speed-volume relations and the nature of the existing road network, the increase in congestion will be greatest near the work place. As shown, the area between rr and q'q' representing the true costs of congestion resulting from added population could well be greater than the area between rr and qq used in the examples.

As another consideration, it was assumed in this section that persons moving to a city were charged for the extra road expenditures due to their entry. Real methods of financing probably come closer to charging an average rather than a marginal cost. In the absence of transfers between cities, which only accentuates the conclusion below if the transfers are due to federal subsidies to urban expressways, financing of roads which breaks even within the city represents average cost financing. If there are diseconomies of scale to additional road expenditures, people in the city are being charged for roads less than the marginal road cost imposed by an additional person. The congestion situation depicted in Figure 3b is suggestive of reasons for diseconomies of scale. When new entrants come to a city, the road system is expanded to accommodate greater travel distance as the city expands, and in addition road expenditures are undertaken near places of work to reduce added congestion there. If the road capacity for greater travel distances from the edge of the city could be provided at constant cost, the added expenditures to reduce congestion at the city center would imply an increasing cost situation considering the road system as a whole. Since density decreases at greater distances from the center, diseconomies of scale are probably eventually reached even in providing for the greater travel distances. A hypothesis for further consideration is that new residents of large cities pay less than the marginal costs of expanding the road

system, further increasing the negative externalities connected with roads above those estimated in this section.

Pollution

Air pollution, leading to external effects which emitters do not take into account, has received attention as one of the chief reasons why social cost may lie above private cost in Figure 1. There is a tendency for air pollution concentrations to increase with city size, although the relationship with city size is far from exact. One expects to find an observed tendency for concentrations to increase with city size, although the relationship with city size is far from exact. One expects to find an observed tendency for concentrations to increase with city size, for all three major sources of air pollution. The three sources are, first, industrial establishments emitting gases becauses of fuel use and because of the chemical reactions connected with particular manufacturing processes, second, household heating and, third, vehicle emissions. Pollution, as reflected in its measurement as physical amount of substance per volume of air, is not a matter of absolute amount but of amount relative to air being diluted. The growth of a city brings more and more emissions relative to the air available for dilution in a vicinity.

The air being polluted is an infinite ocean whose molecules are in motion. Pollutants per unit of time are released into the air and begin to diffuse, so that concentrations of pollutants are highest near a pollution source and diminish as one moves away from it. There are many pollution sources, separated geographically, within any city. One effect of city growth is to lead to increased emissions from existing sources, such as in the industrial parts of the city, at the electricity generating plants and on the streets and highways where traffic volume is increased. Another effect of city growth is to lead to an increase in the number of sources, particularly from new houses and new roads removed from the center of the city. The pollution from the new sources spreads to the nearby existing sources, thereby increasing pollution concentrations in spite of the fact that new sources are somewhat removed from existing sources. If city growth is sufficient to have an effect on residential and industrial density, yet another effect of the growth will be to lead to greater crowding together of sources of pollution. Because there will then be less opportunity for dilution between sources, pollution concentrations in the city will increase.

Cities of a given population vary in amount of pollution for a number of reasons. Types of industry, types of fuel, degree of reliance on automobile transportation and weather differences are among the reasons. These reasons for differences may be compounded by complex interactions between types of pollutants particularly when smog is created. These reasons for variation in amount of pollution other than city population complicate the task of interpreting empirical evidence on pollution and city size, and they also complicate the task of constructing analytical pollution models.

Most of the scientific concern with how air quality is affected by pollution has been devoted to physical models for one or the other type of polluting phenomenon, with applications limited to particular metropolitan areas. Combining urban theory with physical principles from these models, it might be possible to construct simple analytical models indicating how pollution can be expected to vary with city size. The models would need to indicate how pollution concentrations grow at particular points taking account of lack of uniformity of pollution across the city.

Such models do not yet exist. The only measures comprehensively available to compare pollution levels among many cities are readings at a single downtown point. The increase in pollution at this downtown point as the city grows underestimates the increase in pollution nearer pollution sources, and it overestimates the increase for areas in the city not much affected by increases in emissions. If pollution increases most near existing sources, then the downtown measure underestimates the pollution increase where the pollution does the most harm, as will be brought out below. Based on downtown readings, a judgment expressed in a previous paper[1] was that an extra worker added to a city increases suspended particulates by .0001 micrograms per cubic meter. Let ds_i/dn refer to the increase in pollution concentration at the i^{th} grid in the city. If $d\hat{s}/dn$ is the increase in downtown pollution reading per unit increase in employment, an approximation is that this increase is applicable to all grids, or $ds_i/dn=d\hat{s}/dn=.0001$ for all i. Error from this approximation will be considered later.

To find effects on national income, the physical pollution changes considered so far must be translated into costs. The costs of pollution consist of the sum of (1) damages caused plus (2) the defensive expenditures undertaken to avoid damages. Those affected by pollution make choices as between allowing damages to occur and undertaking defensive expenditures to prevent the damages, thus determining how costs are split between the two components. For materials, the costs of more rapid corrosion are an example of damages. Use of more resistant paints and more frequent painting are examples of defensive expenditures. The unhappiness caused by dirty homes, clothes and faces are examples of damages, whereas the extra cleaning, laundering and bathing are examples of defensive expenditures. The discomfort due to tears and unpleasant smells caused by pollution are examples of damages, whereas climate control devices provide examples of defensive expenditures. Suffering due to ill health, work days lost or impaired in efficiency and years of life lost are examples of damages. Medical measures connected with disease caused by air pollution may be viewed as a combination of damages and defensive expenditures undertaken to avoid further damage. Movement of an industrial plant to a location in another city which would be a higher cost location in the absence of pollution, and spreading out of residences within a city to avoid pollution, are examples of defensive expenditures taking the form of locational changes.

---

[1] Tolley, G. S., "Population Distribution Policy," op.cit. A more exhaustive review of evidence on air pollution is contained in a study by Irving Hoch at Resources for the Future, Inc., not yet published.

Figure 4 depicts the costs. Each dashed line shows damages if defensive measures did not vary with level of pollution. Because of the possibility of reducing the deleterious effects of pollution through defensive expenditures, higher concentrations lead to greater defensive expenditures so that there is movement to higher dashed curves at higher pollution levels. The dots indicate the locus of actual damages given the increasing defensive expenditures. To these actual damages are added the amount of the defensive expenditures undertaken at each pollution level, to arrive at the thick solid line showing the costs of pollution. The costs of pollution rise increasingly with pollution level, although the rise is not as great as it would be if there were no defensive expenditures. Notwithstanding the defensive expenditures, at a sufficiently high pollution level the costs would rise practically vertically as human life became sustainable only with gas masks.

Let g_i be the increase in cost to a family resulting from a one unit increment in air pollution in the i^{th} grid, as determined by the slope of the thick schedule of actual damages plus actual expense of defensive measures shown in Figure 4. For the reasons brought out, g_i will become progressively larger at higher pollution concentrations. An approximation is that the increase in costs resulting from an increment in pollution has the same estimated value $\hat{g}$ in all grids in the city of $g_i = \hat{g}$ for all i. Based on analyses of property values, an estimate of g from an earlier study is that an increment of one microgram of suspended particles per cubic meter increases cost to a family by $2.40 to $4.50 per year, or by an average value of $3.60 per year.

The information on costs must be combined with the information considered earlier on how pollution is affected by city growth. By multiplication, the increase in costs of air pollution in the i^{th} grid resulting from an expansion of city employment by one worker is equal to the number of families in the grid n_i times the increase g_i in costs resulting from an increment to air pollution times the increment to air pollution ds_i/dn caused by the expansion of employment, giving a cost for the grid of $n_i g_i (ds_i/dn)$. For the entire city the cost is the summation over all grids, giving as the true cost of adding a worker $\Sigma n_i g_i (ds_i/dn)$.

Using the approximations noted above, it can be seen that the estimated increase in air pollution cost from adding a worker to the city reduces simply to $N\hat{g}(d\hat{s}/dn)$ where N is the total number of families in the city. Insertion of the coefficient values that have been noted for $\hat{g}$ and $d\hat{s}/dn$, for a city of 6 millions persons having 1.5 million one-worker families, gives the estimated cost of adding a worker as 1.5 million times $3.60 times .0001, or $540. If eliminating a 10 percent growth in population would prevent the addition of 150,000 workers to the city, the saving in air pollution costs would be $540 times 150,000 which equals $81 million.

The difference per family between the true cost and the estimated cost is $\Sigma(n_i/N)g_i(ds_i/dn) - \hat{g}(d\hat{s}/dn)$, which is obtained by subtracting the esti-

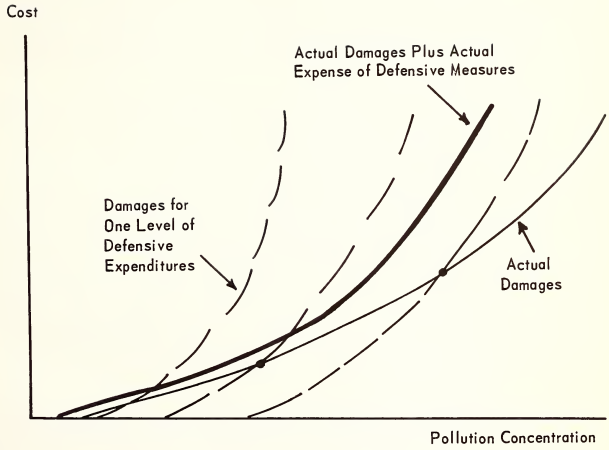


FIGURE 4. Air Pollution Cost

mated cost from the true cost and dividing by the number of families. If all grids experienced the same increase in pollution as at the downtown reading and if the cost of additional pollution were the same in all grids, then clearly the estimated cost and the true pollution cost would coincide, i.e. the foregoing difference would be zero. Suppose half the grids receive no increase in pollution, while the other half near polluting sources receive an increase in pollution 2 times that at the downtown reading. Suppose further that the half receiving the increase in pollution, being already heavily polluted, are at a steeper slope on the thick schedule showing pollution costs in Figure 4 than the average of cases to which the estimated slope $\hat{\beta}$ applies. The supposition that the slope for the heavily polluted areas is 5 times $\hat{\beta}$ provides a liberal allowance for error in slope, since observations on heavily polluted areas themselves enter into the statistical estimate of $\hat{\beta}$. Under these suppositions, the true pollution cost for the heavily polluted areas is 2 times 5, or 10, times the estimated cost per residence. The true cost considering the city as a whole is thus an average of half the observations at zero and half at 10 times the estimated cost, making the true cost for the city 5 times the estimated cost.

Another reason the foregoing estimate could understate true cost is lack of knowledge of harmful effects of pollution on the part of those being harmed. The early study reported results from Lave and Seskin and suggested that a microgram of suspended particulates may increase the death rate by 4 persons per million. Using the estimate above that a worker adds .0001 micrograms, the addition of a worker then increases the death rate by .0004 for a city of 6 million persons. A liberal allowance is that the unknown consequences of loss of life and sickness from diseases caused by air polution are $225,000 per life lost. This appears liberal because there is in fact some awareness of the deleterious effects of air pollution on health. The $225,000 figure is that which would just make the unknown costs equal to the known costs as estimated based on land values. That is, $225,000 is a $540 cost per worker divided by the increase in deaths of .0024. Put another way, it would take an unknown cost per life lost of $225,000 to make the total costs of air pollution double what they were estimated to be on the basis of the perceived costs estimated from land values.

The estimation has ignored spatial adaptation to pollution. If all pollution took place in completely automated industrial parks having no effect on residential areas which themselves emitted no pollutants, then additional growth of a city would need cause no increase in pollution costs. The conditions are clearly not fulfilled. Yet there is some tendency toward segregation of heavy pollution to areas where effects on others are reduced. To avoid pollution and other disamenities, residences tend to locate away from areas of heavy industrial pollution, which however may result in greater commuting costs. Because of land available, economies of association and zoning, industries tend to locate near one another and so have their main pollution effects only on the workers there and on the relatively few residences located nearby. For industrial pollution, therefore, there is a possibility costs are significantly lowered by spatial adaptation. The reasoning does

not suggest the costs are completely eliminated.

For residential and automotive pollution, those bearing pollution costs as a group tend to be those emitting the pollutants. There appear to be only minor possibilities for them to escape their own pollution through spatial adaptation.

For all types of pollution, more stringent environmental controls imply there will be less additional pollution with future city growth than there has been with past growth. In sum, the original $81 million estimate of cost savings from keeping a city of 6 million from expanding by 10 per cent could be an underestimate in view of the greater costs of pollution in areas already polluted and in view of the unperceived costs by those harmed. The reasoning suggests methods for allowing for error. The estimates of error, while largely illustrative, suggest the possibility of obtaining an idea of the maximum possible costs. In the illustration given, multiplying the maximum estimate of greater costs of pollution in areas already polluted using a factor of 5, times the maximum effect of unperceived costs using a factor of 2, gives a maximum possible cost of 10 times the original estimate, or $810 million. Not quantified were neglected effects reducing costs, namely, spatial adaptation to pollution and lowering of emittants due to environmental controls.

Implications

The examples in this paper contribute to the discussion of national growth policy by providing indications which have been lacking about the magnitude of urban environmental externalities. Most of the examples in the paper concerned the effects of retarding the growth of a city of 6 million persons by 10 per cent or by about 150,000 workers. While the numbers are only illustrative, they are realistic. The private loss from not allowing the city to grow, as inferred from extra commuting it would be worthwhile to undertake, was estimated as $43 million. Assume that the savings in congestion costs are as in the first congestion example or $47 million. Later examples gave lower estimates, but reasons were noted for underestimation in all the numerical examples for congestion. Assume that the saving in air pollution costs is $81 million as in the original pollution example. A maximum possible estimate was developed of 10 times this amount, but this maximum would be reduced by nonquantified factors making costs lower than estimated. The sum of the savings in pollution and congestion costs, less the private loss, gives a net gain of $85 million. Density effects are not included on the assumption that the city in question is too large for its density to be affected by the peripheral growth.

For comparison, wage payments provide an indication of the total contribution to national income of the 150,000 workers who would be excluded from the city. Wage payments of a little over $6000 to each worker would total to $1 billion. This is an estimate of how much national income would be produced by the workers if there were no environmental effects. The gains

are then 8.5 percent of the total product of the workers whose location would be affected by the policy. On the other hand, compared to the total product of workers including those who would remain in the city as well as the 150,000 who would be excluded, the gains are only .85 per cent of total product. These are large amounts, but failure to realize them is clearly not calamitous. The results suggest that environmental effects should be an important, but not overwhelming, consideration in population distribution policy.

For the gains to be realized requires that there be no losses in the alternative locations where the workers kept out of the city would be employed. According to these examples, the gains could be realized by scattering the workers in centers sufficiently small not to have environmental costs.

What are effects of population redistribution not considered in this paper? Noise, urban blight, water pollution and crime may be mentioned as additional environmental costs increasing with city size.[1]

Quantifying these additional environmental effects in dollar terms would almost surely increase the estimated gains from the policy considered here. Nonenvironmental effects might increase or decrease the gains. Nonenvironmental effects include economies of scale, which may affect the slope of the private city cost curve and may also have external effects. Unemployment and underemployment can lead to net income effects, as often mentioned as occasionally estimated. The most important reason for effects of population redistribution could be institutional externalities. These are external or unperceived effects of changing location due to differences between costs paid and benefits received in the public decision making process. Differences between payments and benefits from road services arose as an institutional externality in this paper. The supposed attraction of poor people to cities with high welfare payments has received publicity, but with little in the way of statistical support. In view of the importance of education costs as a factor in local finance, differences in payments and benefits from education particularly deserve attention.

Most environmental effects fall under the general heading of national income effects, in that they lead to changes in well being obtained from goods and services produced by people in everyday activity. The same is true of the other effects mentioned in the preceding paragraph. There remain effects of population redistribution that, even under the broadest definition, are not part of national income. These include goals in the distribution of income and in the integration of different income and ethnic groups.

Whether the environmental gains estimated in this paper would be offset by the various other effects of population redistribution can be ascertained in future work if there is commitment to seriously estimate the magnitude of those effects.

---

[1] The Hoch study referred to in a previous footnote reviews evidence on the association between city size and a number of environmental factors.

ESTIMATING THE REGIONAL IMPACT OF FEDERAL INCOME MAINTENANCE PROGRAMS

T. PAUL SCHULTZ

Much of the effect of the Federal Government on regional development is inadvertent. Overriding interests dictate the timing and direction of change in public policy which are often the source of substantial regional repercussions. These regional consequences of major pending legislation should, therefore, be appraised by those that are engaged in the planning and coordination of development policies for depressed regions.

An important piece of social legislation now pending before Congress would change substantially the system of rewards and taxation that bear upon the working and nonworking poor. Current welfare programs differ greatly in their generosity across States, and most State programs exclude the working poor. Therefore, the proposed Federal legislation is likely to benefit a large population of working poor, radically changing their work incentives, and to increase current levels of income maintenance in most of the less affluent States, particularly those in the South. This shift in support levels and work incentives will not have a uniform effect on all regions, because the poor are disproportionately concentrated in urban ghettos and Southern rural areas, and the increases in support levels will occur predominantly in those regions where rural poverty is highly concentrated. Our research task is to propose an analytical and statistical framework to assess these differential regional economic and demographic effects of alternative Federal income maintenance programs (IMPs).

Rand's research will stress the development of a methodology that can be applied to available U.S. data to draw inferences on the primary effects of various IMPs on interregional migration and labor supply responses of households. Less attention will therefore be given to the complex problems of translating these primary effects into shifts in consumption patterns and secondary and tertiary regional consequences. Much uncertainty exists as to how precisely households in different regions will respond to different IMPs. Integrated applications of emerging economic theory of household behavior with modern econometric methods to newly available household surveys and the 1970 Census Public Use Samples should permit much improved predictions of these household response patterns. The problem of estimating the secondary consequences for regions of migration and labor supply changes have been studied elsewhere, and will therefore not be stressed in the Rand project.

To evaluate the effects of an income maintenance program it is necessary first to define the feasible set of alternative forms that much legislation is likely to take. With this in mind, a review as undertaken of various proposed legislation considered by Congress to isolate the range of critical

parameters and characteristics of politically acceptable programs. In large part, the program differences are one of degree, not direction, but there are a number of new features in each proposal -- for example: (1) the treatment of Social Security and personal taxes; (2) the list of exemptions and restrictions for the computation of taxable income; (3) the income support level; (4) the implicit tax rates on earnings; (5) the economic and demographic characteristics of the beneficiary population. These differences may require slight modifications of the specification of the estimation equations and the subsequent computer simulation to evaluate the regional impact of the program, but these differences need not modify the general approach.

To obtain estimates of the impact of an IMP on a region's labor supply and composition, analysis must encompass a wide range of interrelated areas of behavior within the household. A disaggregated approach is also required, for frequently the susceptibility to migration differs by age, sex, family status, schooling, race, and material assets. The response of labor supply to similar conditions differs across these same characteristics in the population. Response parameters must therefore be estimated separately for many of these economic-demographic groups in the population. These response estimates in conjunction with the compositional breakdown of the population by region and income level obtained from, say the 1970 Census, then permits one to estimate the primary regional impact of a particular program regime on labor supply, wages, and transfer payments. Implicit in these estimates would be the program's effects on the composition of a region's labor supply. Precisely how these changes in personal disposable market income and nonmarket time are used may require the estimation of activity, expenditure, and savings functions for the demographic groups of IMP beneficiaries.

Our general analytical model of household decision making views the family as combining resources at its disposal -- the time of family members and goods and services purchased in the market -- to satisfy its wants. The family decides how much time and whose time to allocate to market and nonmarket activities on the basis of the values of time of the family members in these two sets of activities and the nonemployment income of the family. It follows that family members specialize in either labor market or nonmarket activities according to their comparative advantages. Because an IMP taxes the market wage but not the return to nonmarket production, workers receiving income maintenance payments will tend to withdraw time from the labor market and reallocate it to nonmarket activities. Conversely, families that currently receive welfare under which provisions earnings are fully offset by a reduction in welfare benefits will have the incentive under the new uniform income maintenance program to spend less time in nonmarket activities and allocate more time to the labor market. The regional impact of a welfare reform on the supply of labor will therefore depend critically on the proportions of beneficiaries in these two groups -- working poor not currently on welfare, and current beneficiaries of the welfare system.

This reallocation of time between market and nonmarket activities will change the household's money income and its supply of goods and services from the market, influencing also the productivity of time spent in nonmarket activities. These interrelationships within the household that underlie the market supply of labor have not been analyzed satisfactorily; yet they appear to be at the very core of any policy implications drawn from all prior studies of the effects of welfare reform on the supply of labor.

MODELS OF LABOR SUPPLY

The deficiencies of labor supply studies can be divided into those associated with the general conceptual model, with the statistical estimation techniques, and with the appropriate use of empirical material.

Income maintenance programs affect the household's allocation of resources in two distinct ways. The substitution of time toward activities that are less taxed by the IMP is a logical consequence of, and the inefficient distortion from, a specific tax, holding constant all available resources. This pure price or substitution effect of an IMP on the incentives to work in the labor force will be most noticeable for individuals who have a wide range of attractive opportunities for nonmarket work and who can specialize in these nonmarket activities because other members of their household are engaged in the labor force. The substitution effects of an IMP on labor supply are thus likely to be relatively less for unmarried heads of households than for unmarried persons residing in (but not heading) a household. This should be generally true for adults who have never married or are not currently married. For similar reasons among currently married individuals whose spouse is present, the extent of market and nonmarket specification afforded by the family tends to increase the elasticity of market labor supply for the wife and decrease it for the husband. The anticipated longevity of the marriage, the couple's tastes for market and nonmarket goods, the spouses' comparative advantages in nonmarket work, and lastly, the market after-tax wage opportunities for them will determine how they form, over their lifetime together, special skills in market and nonmarket activities. Cross substitution effects between the time of spouses have been generally ignored in empirical studies of labor supply. We plan to estimate their importance among the working poor.

An income maintenance program also augments the total sum of resources available to beneficiaries, who may use this increase to purchase market goods, produce nonmarket goods, and use their time to consume both. It is traditionally assumed that since "leisure," an abstract form of pure consumption, is thought to be a superior good, increases in income will elicit greater consumption of leisure and consequently less labor will be supplied to the market. When, more realistically, time can also be allocated to nonmarket production, there is less reason to assume that "leisure" and nonmarket activities are both superior (relative to market goods), and hence less reason to conclude that increased income will consistently curtail the supply of labor to the market.

Although economic logic does not prescribe the direction of the income or wealth effect on market labor supply, there is a long history of econometric searches for a plausible positive income effect associated with nonemployment sources of income. As noted before, single person households are more constrained in their opportunities for market and nonmarket specialization, and therefore might be expected to exhibit a more pronounced reduction in their market labor supply as their nonemployment income increased. Among couples who have developed their own comparative advantages in market and nonmarket work, it is unclear a priori whose market supply of labor would be more responsive to an increase in nonemployment income.

The indeterminancy in sign of the income effect of an IMP on the market labor supply of beneficiaries is only the start of the analytical problems associated with the meaning and measurement of income effects. A variety of decisions persons make have a simultaneous bearing on their present and future lives together, and they undoubtedly influence the process of life cycle human and nonhuman wealth formation. Since a large share of nonemployment income is directly related to past savings and investment decisions, which were influenced by human capital formation and life cycle earnings opportunities, the current flow of nonemployment income bears the imprint of past choices to work and save which were influenced by past tastes for work, time preferences, etc. These past influences on work and savings behavior are very likely to have a continuing influence on the individual's current behavior.

In circumstances such as these where an explanatory variable, in this case nonemployment income, contains the unexplained effects of factors that are thought to influence independently the dependent variable (labor supply), partial associations among the variables need not provide a consistent basis for estimating the nature or magnitude of the income effect. For parallel reasons, the market wage may also provide a biased basis for directly estimating substitution effects on labor supply. In both instances simultaneous equations techniques are required to obtain consistent estimates of the regional impact of an IMP on labor supply. Our research will elaborate and refine this new approach for the estimation of income and substitution effects on labor supply.

In addition to the general equilibrium conceptual framework within which the market labor supply decision will be analyzed, and the simultaneous equations techniques proposed for estimating the income and substitution effects of an IMP, the Rand research will explore the implications of the choice of sample for analysis and the appropriate measure of labor supply and wages. To a large extent we have sought to include all potential IMP participants in our working samples, and we will only at a later stage examine the stability of estimates over alternative, more restricted, samples. The adoption of simultaneous equations techniques resolves several problems related to the measurement of wages and nonemployment income for the poor, but it also raises new problems concerning the comparability of wage opportunities for persons in and out of the market labor force.

In summary, the Rand approach to estimating the labor supply effects of a Federal income maintenance program synthesizes a number of new directions for research that have only begun to crystallize and have not yet been applied in an integrated manner to this important policy question. The powerful concepts of household production in both the market and nonmarket activities will play a central organizing role. The use of simultaneous equations estimation techniques that are designed to explicitly cope with interrelated economic and demographic behavior of persons and families over their life cycle should also permit us to limit definitional biases and problems related to unobserved tastes for work, wealth, and life styles that have marred earlier studies.

PRELIMINARY EMPIRICAL FINDINGS

I have discussed, this far, only the elements of the analytical and statistical framework needed to assess the regional effects of welfare reform. To document that the shortcomings of earlier studies of labor supply behavior are more than academic, we have analyzed several bodies of data and estimated labor supply functions by the methods used in the past and by those methods we are proposing. Although highly tentative at this stage in our work, let me cite some empirical implications of these exploratory investigations.

We have found, as anticipated, that the secondary worker who is not the head of their own household is very responsive to changes in their current after-tax wage rate. Heads of households are less responsive, perhaps for the reasons discussed above. Married white women with husband present also respond to the wage rate of their husband, but this is not as noticeable for black married women. Black married women supply labor to the market in much the same manner as a household head, even though the husband may be present. All of the labor supply response parameters for secondary workers are increased substantially by using the two-stage estimation technique of imputing a "permanent" wage to each potential worker rather than using observed market wages to estimate parameters directly from the segment of the population currently in the labor force.

The income effects are not stable or substantial across estimation techniques and they are also sensitive to the concept of nonemployment income adopted in the analysis. These results justify our focus on the wage-tax effects, as the crucial dimension of welfare reform in terms of its impact on the supply of market labor. If these preliminary results are confirmed in a more thorough empirical study planned for next year, they suggest that a substantial reduction in the labor supply of secondary workers would occur in the South, especially in rural regions, with the adoption of welfare reforms. On the other hand, in the urban North, where a larger fraction of the poor are already on welfare encountering confiscatory taxes on earnings, the effects on labor supply of welfare reform might be less pronounced and distributed more evenly between primary and secondary workers.

A MIGRATION MODEL OF FAMILY CHOICE

In conclusion, let me say a few words about our preliminary approach to migration, which has been the responsibility of Julie DaVanzo. To understand how an IMP will influence the migration process requires essentially a new conceptual start, because the empirical and theoretical literature on migration is not as advanced as that dealing with labor force participation and supply. Historically migration models have progressed from mechanical views of the world (e.g., gravity models) toward more behavioristic models (e.g., human capital). We will proceed in the latter tradition but will depart from previous work in three major ways: (1) from the human capital framework we will emphasize both the causes for differences in the rate of return to migration (demand) and the causes for differences in the personal cost of investable funds required to finance migration (supply); (2) the decisionmaking unit is the family rather than the individual for married individuals, and hence the woman's employment opportunities enter the male's migration decision and vice versa; and (3) a broader spectrum of regional attributes than earnings are considered that might influence the migration process.

Initial empirical work guided by this model has examined census interdivisional gross migration flows for 1955-1960 by sex, age, race and schooling. It is already clear that this model outperforms others, because it considers the woman's employment opportunities and integrates these into the family decision making process. Improved predictions of gross migration flows have been obtained for men and women. Origin income, as a proxy for wealth, appears to assist in the financing of the migration process, accounting for the widely noted tendency for migration to be stifled by excess or transitory poverty. More refined testing of the model is now underway using micro survey data for unrelated individuals and families. Unfortunately the best data sources for wealth and wages are the least satisfactory for the study of migration since they are often limited to one point in time. A number of longitudinal surveys will be examined for their potential to alleviate these problems such as the Rand Social Security Continuous Work History 1% Sample, the Parnes Manpower files, the Michigan Income Dynamics Survey and various 1970 Census Public Use Samples on which migration histories and current residence codes may not have been suppressed.

CONCLUSION

There are many, as yet unanswered, questions regarding the details of Federal measures to reform U.S. welfare policies. Even if these changes are far more modest than initially envisioned by advocates of the Negative Income Tax, they are likely to have important effects on the regional supply and composition of market labor and on the regional distribution of transfer payments. To estimate these effects requires knowledge of how labor supply and interregional migration decisions will respond to the changed system of incentives and taxes. The theoretical and analytical framework being developed in this research project should advance our ability to assess these primary impacts of alternative income maintenance programs on the economic and demographic characteristics of regional populations.

ASSESSING THE DIFFERENTIAL REGIONAL CONSEQUENCES OF FEDERAL TAX-TRANSFER POLICY

STEPHEN P. DRESCH

Whatever the ultimate legislative fate of the Nixon Administration's domestic policy initiatives, it can at least be claimed that Administration sponsorship has brought to the stage of politically serious debate a number of proposals which heretofore had been the subject only of academic discussion. In each of the most noteworthy instances, revenue sharing, welfare reform and restructuring of educational financing in conjunction with a value added tax, the Administration's role thusfar has not been to enact legislation but to focus attention both on the general type of program it has advocated and, perhaps more significantly, on the implication of existing policies and programs in these areas. Thus, the perceived likelihood of a concrete value added tax proposal has been a major source of the current legislative concern with federal income tax reform; revenue sharing has led to a searching look at the vast system of categorical grant programs which has developed almost by accretion over three decades; and the real inequities and inefficiencies of the existing welfare system have been placed in sharp contrast with a Family Assistance Plan which would go far to render at least the federal component of the system broadly progressive.

As will become evident, my purpose here is not to extol the specific proposals of the present Administration, but rather to develop, somewhat prematurely, the most immediately policy-relevant results of my own work, and that of others at the National Bureau of Economic Research and elsewhere, much of it supported by the Office of Economic Research of EDA, and more importantly to suggest the need for particular types of comprehensive analyses of such federal programs.

The analytic perspective which informs this effort, if only vaguely and incompletely at this stage, derives broadly from the field of national income determination, in which economics has been most successful in achieving policy relevance. In the assessment of national economic policy, it has long been recognized that a closed model, or conceptual framework, is essential to meaningful results. That is, it is meaningless to consider the effects of broad-based policies, policies which affect all or a substantial proportion of the economy, without incorporating system responses to the full policy menu. To take a concrete example, it is obviously absurd to say that to double the government's share of GNP all that is necessary is to double government expenditures, or, equally absurd, to consider in isolation the effects on national income of federal expenditure policy, on the one hand, and of federal tax policy, on the other, and then to simply add the effects, in some undefined arithmetic sense, to predict the impact of federal fiscal policy on the level of national income. The point is,

simply, that different policies, and components of policies, interact and that the effects of a policy menu must take into account the consequences of policy on non-governmental behavior.[1]

This point has obviously been learned with respect to national fiscal policy, but in most other areas of federal policy it is an idea which has been very slow in taking hold. Thus, we assume that federal housing subsidies will not affect the price of housing, we ignore the effects of various direct and indirect subsidies to employment expansion in particular areas on employment levels elsewhere, assuming indeed that they have any effect on employment anywhere, and in almost all cases in assessing the effects of a federal program we ignore the fact that the program has costs in terms of alternative programs foregone, and that these costs also have effects which interact with those of the program under examination.

The primary explanation of these inadequacies in public policy analysis is ultimately the complexity of the system with which we are dealing. Even for questions of aggregative fiscal policy, closed, general equilibrium models can become incredibly complex. But assessment of more disaggregated policies and policy consequences involves increases in complexity by many orders of magnitude.

Analytical capabilities for dealing with these issues will not develop instantaneously, but their evolution will be fostered if we at least begin to develop an appropriate conceptual framework, one which explicitly identifies the major elements which should be treated in a closed public policy model. This paper is concerned primarily with one of the basic inputs into a closed model: the specification of the complete policy menu to be subjected to analysis. The first parts of the paper are concerned with intergovernmental grants. Because the participation of the federal government in the existing income maintenance system is primarily through such grants, this component of the grant system is considered first. Attention is then focused on the remainder of the federal grant-in-aid system and on the various restructurings of this system which have been proposed. Changing the focus somewhat drastically, the final section of the paper is concerned with the differential regional consequences of changes in broad-based federal taxes, specifically the substitution of a value-added tax for the corporate income tax. In most cases the assessment of consequences is limited to interregional income redistributions implied by the specified policy menu, with only marginal reference to effects in such areas as interregional growth and income differentials, public service provision, labor and capital movements, changes in the intra-regional income distribution, etc.

---

[1]A more extended discussion of these issues and of their implications for public finance research and policy analysis is presented in John Bossons and Carl S. Shoup, "Analyzing the Effects of Large-Scale Changes in Fiscal Structure: A Proposed Systems Approach," in New Challenges for Economic Research, 49th Annual Report of the National Bureau of Economic Research (1969), and in Carl S. Shoup, Public Expenditures and Taxation, NBER 50th Anniversary Colloquia Series, Economic Research: Retrospect and Prospect, Volume IV (1972; forthcoming).

1. Interregional Implications of Federal Income Maintenance Programs

If contributory and veterans pension programs are excluded, federal programs for income maintenance function entirely through grant assistance to state income maintenance schemes rather than through direct federal transfers to individuals. However, the programs of federal welfare assistance differ significantly from other types of federal assistance to state and local governments, both in intention and in the nature of federal participation. Furthermore, restructuring of the federal income maintenance system will almost necessarily involve changes radically different from those proposed in the areas of other federal grants-in-aid. Therefore, it is most meaningful to examine the implications of these two classes of federal programs separately.

Under both of the major federal welfare assistance programs, Aid to Families with Dependant Children and Aid to the Blind, Aged and Disabled, the determination of eligibility for assistance and of support levels is a prerogative of the states. The only effective federal requirement is that the programs be at least nominally uniform within states. Then, on a formula basis the federal contribution is determined as a function of actual disbursements and average payment per recipient. This federal contribution declines from 100% at relatively low levels of assistance to 50% or less in high assistance states.

Thus, the gross benefit to a state of the federal program is a function of (a) two interrelated state determinations, eligibility and level of support; and of (b) the actual number of recipients. Actually, the state also influences the number of recipients, since the proportion of those who are both legally eligible and receive assistance can be influenced by the administration of the program. Thus, states (legislatures and welfare bureaucracies) determine levels of support, eligibility and the proportion of the eligible population receiving assistance. The pool of loyally eligible potential recipients then serves to finally determine both total assistance payments and the federal contribution.[2]

The resultant geographic pattern of federal contributions is precisely that which would be expected of a program with these characteristics, as indicated in Table I-1. In the case of federal contributions per capita (resident, not recipient) no clear pattern emerges. Obviously, variations in the three factors determining the federal contribution result, to a high degree, in a wash.

However, if the number of low income families, i.e., those with incomes

[2] A formal model embodying this "two stage" approach to state income maintenance policy, almost precisely as indicated here, has come to my attention since this section was written; see Frank A. Sloan, "The Impacts of Grants-in-Aid on State Income Maintenance Decisions," Report No. R-894-FF, RAND Corporation (February 1972).

less than $3,000 is taken as an index of the size of the potential pool of recipients under a standardized national income maintenance program, and if actual federal contributions per poverty family are utilized as an index of the characteristics of state program (effective eligibility and levels of support), the expected pattern emerges: federal contributions are strongl positively related to regional income. Recall, these are not total payments but federal contributions per poverty family. Because of the decline in relative federal contributions at higher levels of assistance, actual assist ance variations are even greater.

This is not to say that income maintenance programs are necessarily more perverse in low income states. Considering the costs to the state (out of own funds) of increases in total support levels and of liberalization of eligibility standards, due to the very large relative poverty populations in low income states, and also considering the relatively low incomes of the non-supported population, total assistance payments, and thus federal contributions, would be anticipated to be positively related to income.

In fact, variation over states in the federal contribution per poverty family (an index of total support levels) is statistically more closely related to the degrees of urbanization than to income, although these characteristics are themselves highly positively related:

$$FC/PF = \underset{}{-2024} + \underset{(1.8)}{399Y} + \underset{(2.4)}{13U} + \underset{(1.6)}{17Pov} + \underset{(0.4)}{189(A/P)}, \quad \bar{R}^2 = 0.38$$

where FC/PF is federal income maintenance contribution per poverty family:

Y is per capita personal income in thousands of dollars;

U is the urban percentage of the state population;

Pov is the percentage of families with incomes less than $3,000;

P is population;

A is area in square miles (A/P is area per person, the reciprocal of population density); and

t-ratios are indicated in parentheses.

Effectively, a 15 percentage point increase in urbanization and a $500 increase in per capita income (each approximately one standard deviation) have virtually equivalent effects on the federal contribution per poverty family, increasing it by $200.

While broad evaluation of the existing welfare program is beyond the scope of this paper, the following observations can be made: On horizontal equity grounds, i.e., equal treatment of equals, the existing system is

obviously a failure, producing interregional variation in federal contributions per poverty family of 400 percent ($235 in the Southeast versus $1,189 on the West Coast). Similarly, to the degree to which welfare-differential induced migrations to urban areas produce substantial external costs (increased burdens on municipal services, exacerbation of housing problems, etc.) a rational system would imply opposite differentials, higher benefits in rural areas to stem migratory flows. On the other hand, it could be argued that such migration is desirable, i.e., that interregional differences in income, reflecting differential factor earnings (employment opportunities) would be reduced through migration, and that such flows would increase total income and output.

However, it is rather obvious, I think, that the weight of evidence is that the present system of federal participation in income maintenance is perverse. This fact becomes particularly striking when the interregional income transfers implied by the system are examined. This aspect of the program can be assessed, of course, only if a method of financing the federal welfare system is stipulated. For present purposes it is assumed that if the federal system of income maintenance contribution were eliminated, income taxes of high income taxpayers (those with adjusted gross incomes in excess of $10,000) would be reduced equivalently via a negative surcharge, or alternatively, that the system is financed by a similar positive surcharge. The implied regional federal tax costs of the $4.1 billion federal welfare program are indicated in Table I-1.

The net interregional redistributions are rather striking. The South Atlantic area, third lowest in terms of income, experiences a net regional outflow, while the South as a whole experiences a net gain of only about $400 million, less than 10 percent of the aggregate federal welfare budget. In contrast, the West Coast (Pacific), the second highest region in terms of income, and of federal tax cost, still reaps a net gain of almost $300 million.

This relatively weak and erratic redistribution to low income areas, particularly the South, contrasts with that which would result from enactment of the Family Assistance Plan. In an EDA-supported study John Kain and Robert Schafer estimate that total federal welfare outlays would increase by about $2.6 billion with enactment of FAP.[3] The South's share of this increase is estimated by Kain and Schafer to be $1.5 billion, while its share of increased federal taxes (or foregone tax reductions), under both Kain and Schafer and my own assumptions, is about $0.5 billion. Thus, the net transfer to the South is about $1 billion, or 40 percent of the increase in federal outlays, in contrast to 10 percent of federal outlays under existing income maintenance programs.

It is precisely at this point that the analysis should not stop. Rather obviously, interregional transfers of these magnitudes would have pronounced

---

[3] John F. Kain and Robert Schafer, "Regional Impacts of the Family Assistance Plan: Some Revised Estimates," Harvard Program on Regional and Urban Economics (December 1971).

regional income effects, i.e., increases in expenditures of transfer recipients would have a multiplier effect on regional income. Indeed, Kain and Schafer estimate that the increase in income in the South would be 2.5 times the net transfer, i.e., $2.5 billion, a 1.3% increase over 1969 personal income in the region.[4] However, thorough assessment of the interregional income effects also requires that the effects of reduction in non-South incomes be incorporated, more generally, that changes in interregional trade flows be specifically identified, and also that feedbacks of the South's employment and income expansion (and the non-South's contraction) on federal taxes and welfare transfers be permitted.

Here, the point can be made very graphically that evaluation of changes in broad based federal policies can not be made on the assumption that "nothing else changes". Federal policies which have impacts on broad sectors and segments of the economy will for that very reason induce responses which must be explicitly considered and which may in fact be the most important consequences of the policies.

To indicate the importance of these indirect effects and feedbacks, the Kain-Schafer study again provides a telling example. Having estimated the initial benefit impacts, increases in incomes of poverty families, they employ a reformulation of Bowles' interregional migration model to predict the effect of FAP on migration behavior.[5] Under their "most plausible" estimates, FAP would reduce net black out-migration from the South only by 2.2 percent while increasing net white in-migration by 3.5 percent. They find these consequences, particularly for Negro migration, somewhat disappointing. However, it should be noted that their predicted migration effects ignore the consequences of the implied interregional transfers on Southern (and non-Southern) income and employment. These consequences would again feed back into the migration process. Specifically, wages and unemployment rates of different categories of labor (skill, race, etc.) would be affected differentially in different regions, resulting in further systematic effects on migration. Thus, the estimated Southern income expansion of $2.5 billion, resulting from the $1 billion net transfer to the South, should have very predictable consequences for black employment and earnings, and hence on migratory flows.

Neither I nor anyone else can at present provide adequate estimates of the effects of FAP in these dimensions. But, I would make one observation: the solution will not be found in interregional input-output or economic base analysis of the conventional type, but rather in a much more disaggregated variant utilizing information on individual firms and establishments

---

[4] Kain and Schafer predict the income effect of the net transfer to the South on the basis of regional exogenous expenditure multipliers estimated by Roger E. Bolton, Defense Purchases and Regional Growth (Brookings Institution, 1966).

[5] Samuel Bowles, "Migration as Investment: Empirical Tests of the Human Investment Approach to Geographical Mobility," Review of Economics and Statistics (November 1970).

in place of many of the presently aggregated industrial sectors. Much of the apparent instability of interregional trade flows will become explicable when individual firms are examined. And in the same vein, locational behavior of new and relocating establishments, which account for possibly 30 percent of all establishments over a two-year period, will also have to be explicitly incorporated.[6] Both our understanding of the determinants of these processes and our capability to handle the vastly expanded quantities of data required are developing sufficiently rapidly to render these issues tractable.

Thus, this rather brief assessment of FAP and of the existing federal welfare system raises more questions than it answers. But, what answers it gives are nonetheless important, for two reasons: it is from these "initial impacts" that further round effects will derive, and only by identifying the initial changes is it possible to determine what ground a thorough analysis should cover.

II. Interregional Implications of Federal Non-Welfare Categorical Grants

Having isolated and disposed of federal income maintenance transfers to state and local governments, this section examines the interregional redistributions implied by the remainder of the existing federal intergovernmental grant system. Because the ultimate objective is to contrast this system with three of the general revenue sharing proposals which have been put forward, the current categorical grant system can, for present purposes, be dealt with in aggregative terms. The discussion will proceed in the following stages: First, what interregional redistributions result from the categorical system? Second, how would this redistributional pattern be altered if any one of the three most prominent general revenue sharing proposals were to be enacted as a replacement for, or as an alternative to increases in existing categorical programs?[7]

The differential regional incidence of the non-welfare-related categorical grants themselves (gross benefit incidence) is indicated in the first column of Table II-1 (per capita grants by region). The degree of variation in per capita grants across regions is substantial, from $43 in the Middle Atlantic area to $104 in the Mountain region, a range of 144% of the lower figure, compared to a national mean of about $56. To

---

[6]Robert A. Leone, The Location of Manufacturing Activity in the New York Metropolitan Area. (Forthcoming: National Bureau of Economic Research, 1972). Of 38,000 manufacturing establishments in 1967, 4,500 had moved within the metropolitan area, 4,000 new establishments had appeared, and 3,000 had disappeared (eliminated or moved from the metropolitan area) by 1969.

[7]This and the following section extend and reinterpret the results of the very preliminary analysis of general revenue sharing presented in my earlier paper, "An Alternative' View of the Nixon Revenue Sharing Program," National Tax Journal (June 1971).

facilitate comparison with the revenue sharing alternatives, it is useful to normalize the categorical distributions to remove the effect of the gross size of the transfer programs. In the second column of Table II-1, the preceding per capita benefits are normalized to a program with mean national grants of $1 per capita. Equivalently, the normalized figures can be interpreted as the ratio of the grants actually received in the region to the grants which would have been received had the grant allocations been made on a purely per capita basis.

In fact, a straight population-based distribution provides one basis for assessing the redistributive implications of the categorical grant programs. Utilizing state indices, 18% of the aggregate categorical grant pool is redistributed from states which lose (by comparison to a population-based distribution) to those which gain. Anticipating the results of the revenue sharing analysis, this relative redistribution will be found to be more than twice as great as that which would be implied by any of the genera formula-grant programs.

Note, however, that the redistribution (again, by comparison to a pure population-based grant) does not seem to be related to such regional characteristics as income or percentage of families with poverty incomes. To assess the relationship between these various characteristics and normalized gross benefits under the categorical aid programs, a simple regression equation was fitted to state observations:

$$CG = 1.69 - 0.19Y - 0.0U - 0.0Pov + 4.68(A/P), \bar{R}^2 = 0.68$$
$$(-0.95) \quad (-0.06) \quad (-0.7) \quad (9.82)$$

The only variable here with any explanatory power is area per capita (the reciprocal of population density); the coefficient of determination between this variable and gross categorical grant benefits is greater ($r^2 = 0.69$) than the adjusted coefficient of determination for the equation. Utilizing only the reciprocal of population density, the equation becomes

$$CG = 0.92 + 4.73(A/P), \bar{R}^2 = 0.69$$
$$(10.44)$$

For a state with average population density (56 persons per square mile), this equation would indicate a gross per capita benefit (normalized) of 1, the national mean. A state with greater than mean area per capita would receive greater per capita benefits, while a state with lesser relative area would receive less.

To comprehend the meaning of this relationship, it is useful to "denormalize" this equation and convert it to aggregate (from per capita) terms:

$$CG' \cdot P = 51.24P + 263.46A,$$

where CG' = 55.70CG: actual categorical grants per capita (national mean = $55.70). Thus, a state receives five times more per square mile than it

does per person. That is, of aggregate federal grants of over $11 billion, the average state receives approximately $222 million, of which about $19 million accrues to the state independently of characteristics other than area, at least of the variables considered here.

This statement is not meant to represent a negative judgment concerning the overall categorical grant programs, simply an observation. However, greater comfort would probably attach if the level of per capita benefits were appropriately related to, e.g., poverty incidence or income. At least in that case it could be argued that federal grants might in the aggregate have appropriate redistributive implications or similar justifications.

One possible justification does stand out. If federal grants are aimed at inducing public investments which are very lumpy and indivisible and have substantial external benefits, i.e., benefits to residents of other states, than large states with small populations might not undertake these investments in the absence of federal aid, and the federal government would be justified in making greater relative contributions in these states for that reason. Even if the externalities are as great or greater in larger (more populous) states, the investments will still be undertaken because in absolute terms the intra-state benefits are sufficient to justify state provision.

The federal highway program has often been justified in these terms. Since this program accounts for almost half of non-welfare categorical grants, it may well explain this observation of an areal-determined federal grant.

Thus far, the discussion of federal categorical programs has been in terms of gross benefits. Assessment of net transfers again requires the specification of a financing mechanism. Again assuming a high-income surtax, the regional grant-financing tax incidence (normalized around a national per capita value of unity) is indicated in the third column of Table II-1. The inter-state variation in the surtax incidence follows very predictable patterns:

$$ST = \underset{}{-1.49} + \underset{(9.53)}{0.64Y} + \underset{(2.67)}{0.0045U} + \underset{(2.36)}{0.0079Pov} - \underset{(-0.33)}{0.05\,(A/P)}, \quad \bar{R}^2 = 0.93$$

Thus, the per capita surtax incidence is closely related to income, to relative urbanization (probably due to special treatments of farm income), and (positively) to the percentage of families with poverty incomes. The latter relationship, which might appear paradoxical in fact is quite reasonable: For any given level of income per capita, the greater the proportion of the population in poverty the more unequal must be the distribution of income, and with a progressive tax, the greater the tax burden.

Net benefits per capita are simply the excess of gross categorical grant disbursements over surtax incidence. The South, which did not benefit

from the gross grant disbursements, experiences a significant net benefit due to its low tax burden. The high income, densely populated Northeast quadrant (New England, Middle Atlantic, and East North Central) bears the full burden of the regional redistributions, providing a net transfer to the rest of the country of $2.072 billion, or 18.6 percent of the $11.1 billion grant pool. The West (Mountain and Pacific), due primarily to relatively low population density, receives almost as large a net inflow ($0.86 billion) as the low income South ($0.94 billion).

Obviously, the predominating actors in the inter-state redistribution are income (through financing incidence) and population density (through the gross grant allocation):

$$CG - ST = \underset{}{3.18} - \underset{(-3.67)}{0.84Y} - \underset{(-0.84)}{0.005U} - \underset{(-1.32)}{0.015Pov} + \underset{(8.83)}{4.73(A/P)}, \quad \bar{R}^2 = 0.78.$$

At the state level the net redistribution due to the categorical grant programs amounts to 30 percent of the total grant pool, i.e., net beneficiary states receive a net inflow of $3.28 billion from the rest of the country. Again anticipating, this degree of redistribution is greater than will be observed for any of the general revenue sharing proposals.

III. Differential Effects of General Revenue Sharing Alternatives

At this writing, it would appear that there is only one politically alive general revenue sharing (GRS) proposal, that recently put forward by the House Ways and Means Committee. However, given the uncertainty which still surrounds the legislative fate of GRS, it does not seem pointless to maintain a broader perspective, specifically to consider not only the "House" variant, but also the proposals of the Administration (the "Nixon" variant) and of Chairman Mills (the "Mills" variant).

Both the House and Mills versions complicate analysis by decomposing the program into two components, one designed for 100 percent pass-through to localities and the other depending very specifically on particular taxation decisions of state governments. Because assessment of the latter requires a rather complex analysis of state responses to the introduction of revenue sharing (apart from some phase-in restrictions which are not particularly interesting in themselves) the present preliminary examination will be restricted to the first components of these proposals.

The major distinctions between the alternatives are: (1) The Nixon variant, unlike the House and Mills variants, would utilize a per capita transfer modified by "fiscal effort" (ratio of own state-local general revenues to personal income); (2) Mills would base the distribution on (a) population adjusted for urbanization (one-half of population plus one-half of urban population) and (b) on the number of poverty families, one-half of the funds to be distributed on each basis; (3) the House variant would allocate one-third of the grants on the basis of each of (a) population (b) urban population, and (c) the inverse of per capita income.

The gross regional distributions under each of the three GRS proposals are indicated in Table III-1, again normalized to $1 per capita nationally. Because of the greater interregional variation in urbanization, poverty incidence and income than in fiscal effort, it would superficially be expected that both the House and Mills variants would exhibit greater variations in per capita benefits than would the Nixon proposal. However, the interesting fact is that the Ways and Means Committee has succeeded in marrying three dissimilar components to produce an almost perfect per capita distribution formula, i.e., the close negative relationship between income and urbanization produces almost perfect offsets between these components, with the result that the House bill differs only marginally from one which utilized only population in the allocation formula. Thus, the House formula, compared to a per capita distribution, redistributes only 1.5 percent of the total grant pool from "losing" to "gaining" states. Twelve states receive allocations within 2 percent of the amounts they would receive under a purely population-based formula.

Thus, the South as a whole receives no disproportionate benefit from House GRS; its relative lack of urbanization offsets the positive effect of its low income. Under Nixon GRS the South actually loses relative to a per capita distribution formula because of its low measured fiscal effort. Here the possible inadequacies of measured fiscal effort appear. Personal income is measured on the basis of residence of the income recipient, while local tax bases, at least in the first instance, include income generated in a particular location, although received elsewhere. Thus, the Mountain and Pacific regions, with disproportionate income of non-resident property owners compared to personal income of residents, to some degree tax that property income and appear to have very high fiscal effort, while the South, with relatively little non-resident property income, appears to have very low fiscal effort.[8]

It is under the Mills variant that the South receives its most substantial relative benefit. Obviously, the factor which predominates is the proportion of families with poverty incomes, and on this basis the region receives 20 percent more under the Mills formula than it would under a per capita distribution.

Although, except for the Nixon proposal, the GRS formulae are based strictly on the variables utilized above to examine categorical grants, it is useful to examine the statistical approximations to the formulae as a preliminary to comparisons with the categorical grant distributions, denoting House, Mills and Nixon GRS by HRS, MRS and NRS, respectively.

$$\underset{}{\text{HRS}} = \underset{}{0.88} - \underset{(11.43)}{0.08Y} + \underset{(26.07)}{0.0047U} + \underset{(5.14)}{0.002\text{Pov}} + \underset{(0.0)}{0.0(A/P)}, \quad \bar{R}^2 = 0.94.$$

[8]For a general discussion of alternative fiscal effort measures, their meanings and inadequacies, see Allen D. Manvel, "Differences in Fiscal Capacity and Effort: Their Significance for a Federal Revenue-Sharing System," National Tax Journal (June 1971).

$$MRS = 0.33 + 0.039Y + 0.0011U + 0.022Pov - 0.0(A/P), \quad \bar{R}^2 = 0.96.$$
$$(1.42) \quad (1.57) \quad (10.05) \quad (0.0)$$

$$NRS = 0.70 + 0.06Y + 0.0U + 0.0Pov + 0.28(A/P), \quad \bar{R}^2 = -0.05.$$
$$(0.50) \quad (0.2) \quad (0.6) \quad (1.08)$$

In the case of both HRS and MRS, the estimated relationships are consistent linear approximations of the actual formulae. From the NRS equation it is apparent that there is no systematic relationship between fiscal effort (ratio of own taxes to income) and the various variables employed.

To move from gross distributions to net redistributions, the previous tax-financing assumption could be employed. However, it seems much more likely, politically, that in the absence of a program of general revenue sharing categorical grants of the existing variety would be expanded. Therefore, prior to examining tax financing, a reduction (foregone increase) in categorical grants will be assumed to provide revenue sharing funds. Table III-2 presents the net regional benefits per capita (normalized) resulting from such a substitution of general revenue sharing in any of the three guises for categorical grants.

Clearly, under House and Nixon revenue sharing the major beneficiaries are the densely populated urban states of the Northeast. They benefit primarily from the substitution of a broadly population-based formula for a categorical distribution which discriminates aginst populous states. The most significant negative beneficiaries are the Western states, which lose for the same reason that the Norteast gains. The low income South is, as a whole, relatively unaffected by the substitution of either of these formulae for the existing categorical distribution.

The Mills variant would produce a redistribution of federal grant funds as radical as either Nixon or House revenue sharing, but in a very different direction. Because of the predominance of the poverty incidence component, the South becomes the major beneficiary, while the Northeast continues to benefit due to the peculiarity of the categorical grant distribution and the West loses even more than under the House or Nixon variants.

To systematize these relationships, net normalized per capita benefits at the state level were regressed on the previous state characteristics:

$$HRS - CG = -0.81 + 0.11Y + 0.005U + 0.009Pov - 4.68(A/P), \quad \bar{R}^2 = 0.69$$
$$(0.54) \quad (1.02) \quad (0.89) \quad (-9.88)$$

$$MRS - CG + -1.36 + 0.23Y + 0.04U + 0.029Pov - 4.75(A/P), \quad \bar{R}^2 = 0.70$$
$$(1.11) \quad (0.26) \quad (2.79) \quad (-9.67)$$

$$NRS - CG + -0.99 + 0.25Y + 0.00 + 0.01Pov - 4.39(A/P), \quad \bar{R}^2 = 0.73$$
$$(1.47) \quad (0.23) \quad (1.25) \quad (-11.01)$$

In net terms, i.e., in terms of the differential benefit derived from GRS as a replacement for (an increase in) categorical grants, it is apparent that the nominal pro-urban features of House and Mills revenue sharing virtually disappear. The operative factors are: (a) the positive poverty relationship from the Mills proposal; and most importantly, (b) for all three substitutions the elimination of the areal element of the categorical programs.

To place these results in perspective, it is useful to convert to the aggregate (as opposed to per capita) benefits which would be produced by a diversion of $5 billion from categorical programs to any of the three GRS proposals. These are indicated by region in the last columns of Table III-2. Under both House and Nixon revenue sharing, the greatest aggregate net benefits are received by the three regions of the Northeast: $420 million under HRS and $354 million under NRS, 8.4 percent and 7 percent of the $5 billion national pool respectively. Correspondingly, the two Western regions would lose $459 million and $342 million under House and Nixon GRS, or 9 percent and 7 percent of the aggregate national amount. The effect on the South is marginal in either case: a gain of $44 million (0.8 percent) with HRS and a loss of $4 million (0.2 percent) with NRS.

Effectively, the Mills proposal would increase the loss of the West to $575 million, 11.5 percent of the national $5 billion, and redistribute it primarily in favor of the South, which would experience a gain of $362 million over its share of current categorical grants, 7 percent of the aggregate pool. The Northeast would still benefit, but at the lesser rate of $202 million, 4 percent of the pool.

A major qualification is obviously in order: this discussion has assumed that general revenue sharing is an alternative to an overall expansion of non-welfare categorical grants of the existing type, the distributional pattern of which would be identical to the aggregate distributional pattern of existing categorical programs. However, categorical programs with very different regional implications could certainly be devised. As will be observed in a later section examining the Administration's special revenue sharing proposals, if introduction of SRS-type programs is the alternative to either GRS or expansions of existing categorical programs, then very different regional redistributive patterns are implied. In brief, the specification of the financing mechanism, the alternatives foregone, is as important as the specification of the ostensible distributional formula.

This can be indicated very graphically by examining the redistributive consequences of the three GRS programs assuming income tax surcharge rather than categorical grant financing. Normalized and aggregate ($5 billion program) net benefits by region are presented in Table III-3. Obviously, the predominant effect of the surtax on the net benefit distributions gives the three GRS options a much less differentiated appearance. For example, in each region the sign of the net benefit (plus or minus) is the same for all three of the GRS programs. For individual regions the Nixon and House alternatives show some variation, but at the South-Non-South level they

are virtually identical, showing net transfers to the South of $419 million and $467 million respectively. And in this instance they differ only in magnitude, not in direction from the Mills option, which results in a net inflow to the South of $787 million. Recall that the assumption of a GRS-categorical grant substitution resulted in a virtually zero net transfer to the South under both the House and Nixon revenue sharing.

The Northeast regions and the West both experience very different transfer effects under the surtax than under the categorical grant reduction. With the latter, the Northeast had received significant benefits under all GRS options, while the West had suffered major losses in all three cases. With the surtax the gain to the Northeast becomes a net loss, and the loss in the West is significantly reduced.

At the state level, net transfers are clearly determined by the incidence of the tax and are thus strongly negatively related to income:

$$\underset{\quad}{\text{HRS - ST}} = \underset{(-11.13)}{2.37} - \underset{(0.19)}{0.73Y} + \underset{(-1.86)}{0.0U} - \underset{(0.35)}{0.006Pov} + 0.0(A/P), \quad \bar{R}^2 = 0.94$$

$$\underset{\quad}{\text{MRS - ST}} = \underset{(-7.49)}{1.83} - \underset{(-1.70)}{0.60Y} - \underset{(3.47)}{0.0034U} + \underset{(-0.12)}{0.0139Pov} - 0.0(A/P), \quad \bar{R}^2 = 0.96$$

$$\underset{\quad}{\text{NRS - ST}} = \underset{(-3.77)}{2.19} - \underset{(-1.06)}{0.59Y} - \underset{(-0.62)}{0.004U} - \underset{(0.99)}{0.004Pov} + 0.34(A/P), \quad \bar{R}^2 = 0.74.$$

In brief, if GRS is tax financed, both House and Nixon revenue sharing become highly redistributive, transferring income from high to low income states. This was not true under the assumption of categorical grant financing. The Mills variant, through both the income-related surtax and the poverty incidence formula remains the most redistributive of the three.

Thus, the interregional redistributive argument for revenue sharing of any variety becomes much stronger if revenue sharing is seen as an alternative to further income tax reduction rather than as an alternative to expanded categorical grants. Most simply, this results from the fact that income tax reductions are of relatively slight benefit to low income areas, while these areas' benefits from categorical grant programs are relatively much more substantial.

Obviously, the redistributive dimension is only a partial basis for evaluating alternative revenue sharing proposals, or for evaluating any given revenue sharing option versus, e.g., expanded categorical grants or federal tax reduction. If the only objective were redistribution, then transfers to individuals and based on income would be a far more effective, certain device. Such transfers would necessarily be interregionally redistributive. Thus, other objectives must also come into play, e.g., allocative and developmental consequences of intergovernmental transfers which could not be achieved via transers to individuals. However, other things being equal, and I would argue that with block unrestricted revenue sharing grants that they probably are equal, appropriate interregional transfer

effects would seem to be among the strongest arguments for general revenue sharing.

Specifically, simulations of state-local fiscal responses to revenue sharing employing the full range of reasonably estimated behavioral parameters indicates that grants of the GRS variety would have only marginal impacts of levels of public service provision. And what effects GRS would have would probably not be deemed to significantly affect such phenomena as regional development. The primary effect would be to reduce (rates of increase in) state-local taxes, in amounts between 50 percent and 90 percent of GRSgrants. Thus, if the major effect of GRS is to replace state-local with federal tax sources, progressivity certainly argues for programs which are significantly regionally redistributive. In addition, since one implication of such a scenario would be greater relative reductions in state-local taxes in low income areas, a developmental case could certainly be made for general revenue sharing if it could be shown that tax differentials were significant determinants of industrial location. In fact, such fiscal factors have not been found to be significant in explaining interregional differentials in employment change. Thus, the redistributive argument for intergovernmental transfers is not greatly strengthened by probable developmental consequences.

IV. Special Revenue Sharing[9]

The Administration's 1971 proposals for special revenue sharing (SRS) represent another attempt at large-scale restructuring of federal intergovernmental fiscal relationships. Like general revenue sharing, SRS would distribute federal grants to state and local governments on a prespecified formula basis. The six SRS programs collectively would, as proposed, incorporate more than $11 billion in federal resources, roughly twice the pool generally considered for GRS distribution.

SRS, however, would differ from GRS in a number of possibly important dimensions. First, while GRS may be an alternative to existing categorical grant programs, i.e., in the absence of GRS existing categorical programs would be expanded by about the same magnitudes, SRS is explicitly billed as a replacement for the existing federal grant-in-aid system. Thus, the six SRS programs would incorporate and supercede over 130 different categorical programs.

Secondly, unlike GRS, at least as originally proposed by the Administration, which would attach virtually no restrictions to the use of GRS grants, the SRS programs would restrict recipient governments application of

---

[9] This section draws upon a working paper by my colleague James Hosek, "An Analysis of the Special Revenue Sharing Proposals" (NBER multilith, 1972; revision in process).

these grants in each of six specified areas of endeavor. As outlined in the Administration's proposals of last year, these would include (1) law enforcement ($0.5 billion in federal grants); (2) manpower ($2.0 billion); (3) urban development ($2.1 billion); (4) rural development ($1.1 billion); (5) general and urban transportation ($2.0 and $0.5 billion, respectively); and (6) education ($3.0 billion).

The basic stated objective of the SRS proposals is the elimination of the minute restrictions and planning difficulties often associated with the more narrowly focused categorical grant (CG) programs. A large number of relatively small and specialized grant administrations would be transformed into a small number of less specialized administrations. Matching fund requirements would be eliminated, as would the need for submission of detailed development and project plans.

The effects of these differences between SRS and GRS cannot be determined a priori. As noted, GRS may itself effectively substitute for categorical programs. With respect to limitations on the use of funds under SRS, difficulties arise in assessing the degree to which such restrictions would be effective. For a state spending more on law enforcement than it would receive under this SRS program, it is not clear how a distinction could be made between law enforcement and tax reduction as the state use of the SRS allocation.

In each of the six SRS programs, funds would be distributed to states on the basis of distribution formulae analogous to those employed for GRS. The formulae would differ, however, from program to program. The law formula, for example, is based only on population, a state's award being equal to its share of national population times aggregate national law enforcement grants. Urban development SRS, on the other hand, employs a weighted sum of urban population, deficient housing units and over-crowding to determine a state's share.

The variables employed in the formulae are apparently selected to provide an indication of need, of the size of the beneficiary population, or both. However, the use of certain variables can occasionally be misleading, and the choice and interpretation of variables should be given serious attention. For example, manpower SRS shares are based in part on the number of unemployed. Many permanent differences in unemployment are due to differences in industrial composition. Thus, a greater number of unemployed in one state than in another does not necessarily indicate a greater need for manpower programs.

Beyond the choice of variables, the weighting given to different variables is of crucial importance. However, no explicit rationale appears to have been given for the particular weights employed. As in the case of GRS the only way to evaluate alternative weights is to identify the differential distributions which would be implied.

The full regional impacts of a substitution of the SRS proposals for existi

categorical programs would depend upon (a) differences in the regional distributions of grant funds under the two alternatives, and (b) changes in state-local fiscal actions resulting from the change in grant structure. Of the two, the second is hardest to assess, and the desirability of an SRS-CG substitution, from the federal point of view, hinges crucially on these behavioral responses. However, for present purposes attention is restricted to the changes in regional grant receipts resulting from the substitution.

To assess these distributional changes, several assumptions and adjustments are necessary. First, the SRS legislation contains "hold-harmless" provisions insuring that in the phase-in period no state would receive less under an SRS program than it had received previously under the replaced categorical programs. If the hold-harmless provisions are incorporated, the quasi-SRS distributions are biased toward the existing CG distributions. Since SRS would ultimately constitute a major restructuring of the grant system, it is more enlightening to compare a "mature" SRS system to the existing categorical programs. Therefore, the hold-harmless phase-in restrictions have been ignored.

Secondly, the SRS legislation provides for increases in aggregate funding of each program over the level of funding of the previous categorical programs. However, to identify the purely redistributive consequences of the SRS-CG substitution, it is necessary to remove the effects of the increased size of each SRS program. Here it has simply been assumed that in the case of each SRS program, categorical programs would be expanded to an identical size if SRS were not introduced. Thus, for a given state j the net benefit of substituting SRS of the $i^{\underline{th}}$ variety, e.g., law, manpower, when an amount CG_{ij} was previously received by the state under categorical programs, is defined as

$$NB_{ij} = SRS_{ij} - CG_{ij}\left(\sum_k SRS_{ik} / \sum_k CG_{ik}\right).$$

In brief, the aggregate size of each class of categorical programs is normalized to the proposed size of the corresponding SRS program.

Net benefits by region resulting from the proposed substitution of the SRS programs for their categorical predecessors are indicated in Tables IV-1 (aggregate amounts) and IV-2 (per capita amounts). Because transportation SRS is composed of two distinct components, general transportation and urban mass transportation, these components are examined individually, resulting in seven specific substitutions.

The net benefit distributions graphically indicate a major result of substituting a formula grant for the existing system of discretionary, judgemental grants. Any formula, by its very nature continuous rather than dichotomous, is going to be less discriminating in the allocation of funds. This is not necessarily a criticism, but it is an important observation of fact.

Consider urban and rural community development. The major net negative

beneficiaries under urban SRS are the highly urban areas of the Northeast and North Central regions, not because they are "discriminated against" by SRS but because categorical programs in the urban sphere were pinpointed in these urban-metropolitan regions. Similarly, the largest losses under rural SRS appear in the East South Central region, which has been a major beneficiary of previous rural development funds. The major beneficiaries are the higher income northern agricultural states (North Central).

The same phenomenon is observed in the case of urban transportation. Existing mass transit programs have been concentrated in the most urbanized centers: New England, Middle Atlantic, East North Central and Pacific. As a result, these areas are the major losers under a continuous distribution formula based primarily on urban population.

Such large changes in regional budgets emphasize that the current categorical and proposed SRS systems of aid make substantially different evaluation of state and regional need. Categorical programs were specifically created to serve relatively special needs and collectively, they are sensitive to many specific dimensions of social want. SRS programs collapse these dimensions into a few indices employed in the distribution formulae. If, indeed, the regional benefit distributions implied by SRS are socially desirable, then as indicated by Tables IV-1 and IV-2, the current categorical distribution must embody substantial misallocations.

Yet, it is difficult to assess the desirability of any program without a formal means of identifying and comparing the full range of consequences of the alternative programs. If, because of more stringent restrictions on the use of federal grant funds, the categorical programs have led to provision of services which are socially desirable and would not have been provided otherwise, due, e.g., to externalities, rather than to state-local tax reductions, then the shift from narrowly focused categorical programs to broad formula distributions may have radical effects on public service provision and public investments. If these restrictions have not been effectively operative, then the most radical change may be simply the difference in geographic distributions. In either event, knowledge of the full range of consequences is necessary for a comparative evaluation of the SRS programs and their categorical counterparts.

V. Broad-Based Tax Substitution Involving the Value Added Tax

As indicated earlier, the necessity for a closed-model approach to questions of fiscal policy is thoroughly accepted by both policy makers and economists. However, even in this field performance has greatly lagged recognition. Existing national economic models are seriously deficient in at least two respects. First, they are highly aggregative, assessing the differential effects of alternative policies only on such large, undifferentiated magnitudes as national income, investment and consumption. At best they explicitly identify impacts only on highly aggregated industrial sectors. The household sector, composed of elements differing radically in terms of, e.g., income and occupation, is rarely decomposed at all. As a

result, very important differences between alternative policy menus in effects on subaggregates are never explicitly observed. Thus, for example, two fiscal policy packages which have identical implications for GNP may differ radically in income distributional terms, but these latter differences, which may not be direct or obvious, are not anticipated and hence do not enter the decision calculus.

The second major deficiency is actually a subset of the first. Just as macro-models in general do not display sufficient functional disaggregation, they also ignore spatial variations in the effects of policy. National economic policy is assessed only in terms of its aggregative national implications, while implications for particular areas or regions may vary greatly. For example, a devaluation, an investment tax credit, and an interest rate reduction induced by monetary expansion may have identical effects on the level of national income, but consequences for particular regions would be expected to differ markedly.

Studies of fiscal substitutions, e.g., the replacement of one tax by another, have been fairly advanced with respect to the first area of macro-model deficiency. That is, their primary concern has indeed been with differential economic and distributive effects of alternative policies on relevant subaggregates. As an example of this type of analysis, a recently completed NBER study by An-loh Lin, David Stout and myself examined the short-run effects of a substitution of a consumption-type value added tax (VAT) for the corporate income tax (CIT), utilizing a 100 sector input-output matrix to identify probable price responses to the change in tax structure, by industry and component of final demand.[10] These price adjustments were then employed to project initial investment, international trade and income redistributive consequences of the tax substitution. This section will first briefly outline the results of that study and will then examine some of the probable differential regional effects of this tax substitution.

Under a value added tax of the consumption type, investment and export sales are exempt from tax. Government may be defined as either subject to or exempt from VAT. However, since the VAT on government expenditures is also revenue to the government, the net VAT base consists only of private domestic consumption. Thus, substitution of a VAT for the CIT amounts to the replacement of a partial income tax by a consumption tax. The serious questions concern the response of prices to: (a) the CIT reduction (or removal) and (b) the VAT imposition.

With respect to the price effects of CIT reduction, classical theory implies that capital earnings in the short run are a quasi-rent, i.e., are not a part of variable cost, and hence that initially a change in tax

---

[10] Stephen P. Dresch, An-loh Lin and David Stout, Substituting A Value Added Tax for the Corporate Income Tax: First-Round Price Effects and Their Implications (NBER 1972,; forthcoming. Draft available on request.)

on these earnings will not be translated into a change in price. Against this theorem of orthodox price theory, a number of "administered price" doctrines have been proposed, which argue that firms set prices not to maximize profit (orthodox theory) but to achieve a predetermined net return to capital. In this latter case taxes on capital earnings are interpreted as a component of cost, and tax changes are thus assumed to be translated into price changes. To avoid becoming mired in this presently intractable debate, the analysis assumes a range of alternatives, from zero "CIT shifting" (the orthodox position) to "full shifting" (the extreme of administered prices). Under the first assumption, VAT-exclusive prices are unchanged by CIT reduction, and net profits rise by the amount of the CIT, while under the second VAT-exclusive prices decline sufficiently to maintain the previous level of net-of-tax profits. VAT- exclusive producer price changes induced by the CIT reduction, by "two-digit" industry, are indicated in Table V-1, for alternative degrees of CIT reduction (S) and shifting (a).

Obviously, the degree of CIT shifting will affect the tax revenue resulting from any given VAT rate. Also, government expenditures will be reduced if VAT-exclusive prices of governmentally-purchased goods and services decline. Appropriate determination of the VAT rate, then, must take into account both of these potential effects of CIT reduction.

For a number of reasons, discussed in detail in the study itself, the VAT rate is determined subject to the condition that the nominal government surplus or deficit be unchanged. Government is defined inclusively, including federal, state and local jurisdictions. Resultant CIT-compensating VAT rates and revenues are given in Table V-2, again for alternative degrees of CIT reduction and shifting.

Having determined the VAT rate, consumer prices can be computed, assuming full forward shifting of the VAT. These VAT-inclusive consumption prices, and aggregate purchaser prices for the VAT-exempt components of final demand, are presented in Table V-3. In assessing the consequences of these price changes, attention is restricted to the extreme cases of full CIT removal with full and zero CIT shifting.

Income Distributional Effects

The tax substitution differentially affects households in different income classes because (a) the proportion of income consumed varies systematically with the level of income, and (b) household wealth holdings similarly vary. With zero CIT shifting after-tax corporate profits rise by the amount of original CIT revenues. In Table V-4 these profit increases are distributed to households on the basis of stock holdings. The increase in nominal consumption expenditure, given the income-specific consumption bundle, is indicated in the second column ("Full Rate") of this table. Net incidence, the excess of the consumption expenditure increase over the share of net profit increases, is also presented. This net incidence declines from 6 percent at incomes of $5000 to $7500 to -7.1 percent at incomes in excess of $15,000. Thus, the tax substitution is significantly regressive.

If the CIT reduction is fully shifted forward through lower prices, then net profits do not increase. However, wealth holders still benefit through the reduction in capital goods prices. These benefits were similarly distributed on the basis of stock holdings, as shown in Table V-5. Increases in consumption expenditures are also smaller in this case, since the decline in prices due to CIT shifting partially offsets price increases due to the VAT imposition. Net incidence is nonetheless regressive, declining from 2 percent at low levels of income to -0.3 percent at the highest levels. In this case, aggregate net incidence is positive because part of the benefit of CIT reduction inures to foreign purchasers of U.S. exports, prices of which have declined.

The effects of fully or partially exempting from VAT those consumption commodities which are income inelastic are also shown in Tables V-4 and V-5. It is readily apparent that commodity exemptions are not an effective means of reducing regressivity, and that such devices as credits against personal income tax liabilities, as proposed by the Administration, should be employed instead. Unfortunately, European VAT and U.S. state sales tax experiences indicate the political strength of the myth that exemptions are effectively progressive.

Obviously, if the tax substitution is regressive nationally, it is likely to be similarly redistributive interregionally. This is indicated in Table V-6, which presents net changes in regional tax liabilities under the two shifting assumptions, distributing the benefit of the CIT removal (profit increase or investment savings) on the basis of dividend receipts and distributing VAT liabilities on the basis of retail sales (a proxy for consumption).

It is readily apparent that low income areas experience significant losses in disposable income. With zero CIT shifting, the increase in tax liability in the three regions of the South reaches $3.05 billion, in excess of 2 percent of disposable income. Full shifting reduces this differential burden to the still substantial figure of $1.2 billion, or slightly less than 1 percent of disposable income.

Here, in fact, we can indicate the relative significance of the regional effects of national policy changes. With respect to federal intergovernmental grant programs, great energy is expended modifying formulae to achieve particular distributions and debating the relative desirability of alternative distributions. However, the net regional redistributions through these programs are almost invariably less than those implied by the ostensibly "non-regional" substitution of a VAT for the CIT. For example, the net increase in the South's tax liability due to the VAT-CIT substitution, between $1.2 billion and $3 billion depending on CIT shifting, exceeds the net benefit to the South of the $4 billion federal welfare grant program ($0.403 billion net Southern benefit), of the $11 billion federal non-welfare categorical programs ($0.939 billion), or of any of three $5 billion income-tax-financed general revenue sharing programs ($0.467 billion for the House program, $0.78 billion under Mills GRS and $0.419 billion under Nixon GRS).

Thus, while we agonize over the regional implications of various grant-in-aid programs, their effects are literally swamped by changes in national tax policy, the regional implications of which are rarely even explicitly considered.

Investment Effects

Because the CIT removal serves (a) to increase after-tax profits and/or (b) to reduce prices of investment goods, rates of return to capital will be increased. Thus, investment should be stimulated by the tax substitution. To assess the probable magnitude of this investment expansion, a liquidity theory of investment was employed.[11] Effectively, the liquidity theory views the rate of gross investment as a function of real net cash flow, defined as nominal cash flow net of taxes and dividends, deflated by capital goods prices.

With zero CIT shifting capital goods prices are unaffected by the CIT reduction. However, after-tax profits increase by the initial CIT liability, as indicated by manufacturing industry in Table V - 6. The last column indicates the probable first year impact of the tax substitution, by the percentage increase in gross investment. This expansion ranges from 2.2% in iron and steel to 22.6% in non-electrical machinery, with a mean of 11.5% for all manufacturing industries.

With full CIT shifting nominal profits (and net cash flows) are unaltered by the tax substitution. However, investment goods prices decline. The change in capital goods prices, determined for each purchaser industry by weighting producer industry price changes by inter-industry capital flows, is indicated in the second column of Table V - 8. The resultant increase in real cash flow and the implied cumulative and first year increases in investment demand are also presented. In this case the short-run investment responses, while still significant, are much smaller, ranging from 4.7% in paper products to 0.7% in chemicals, with a manufacturing mean of 1.5%.

It should be noted that there is a significant difference, apart from magnitude, between the investment expansions in the two CIT shifting cases. With zero CIT shifting the full initial investment response derives from the increase in corporate profits. equal to original CIT liabilities. In this case, the investment expansion is restricted to a corporate sector. Investment in the non-corporate sector is unaffected by the tax change, at least initially. This fact explains the great difference between the manufacturing increase, 11 5%, and the non-manufacturing investment response of only 5.4%. The relatively smaller degree of incorporation outside of manufacturing reduces the overall rate of investment expansion. With full shifting, on the other hand, the investment effect results from capital goods

[11] All investment effects are based on estimates contained in J. R. Meyer and R. R. Glauber, Investment Decisions, Economic Forecasting, and Public Policy (Harvard 1964).

price reductions, which are identical for both incorporated and unincorporated sectors. As a result the total (all industry) and manufacturing investment response are found to be equal, 1.5%.

Once again, the analysis has been in terms of national responses disaggregated only by industry. To estimate changes in investment by region it is necessary to distribute each industry's investment expansion spatially, not a simple problem. Information is available on the geographic distribution of gross plant and equipment expenditure by industry in 1969, but there is no reason why the investment expansion should be distributed proportionately to base levels of investment. Specifically, the increase in investment is much more likely to be "new" investment, i.e., in new plants and in basic new equipment, than is pre-expansion investment. The latter almost undoubtedly heavily reflects the maintenance of the pre-existing spatially distributioned industy capital stock, while marginal investments reflect the changing geographic distribution, a changing distribution which these investments in fact bring about.

In NBER studies just initiated these issues are being directly examined. In the interim it is useful to employ even ad hoc assumptions to reach admittedly tentative conclusions regarding regional implications. For present purposes, it is simply assumed that for each industy any region's share of the investment expansion is equal to its share of base plant and equipment expenditures. However, the fact, for example, that the South's share of plant and equipment expenditures (29.8%) is greater than its share of manufacturing value added (23.3%) probably reflects the differential growth of the South, and thus, simple proportionality will almost undoubtedly understate the South's share of the investment expansion.

If regional distributions for all industries were identical, then under the proportionality assumption each region's share of the investment expansion would equal its share of base investment. However, as shown in Table $\overline{V}$ - 9 and $\overline{\overline{V}}$ - 10 (zero and unitary CIT shifting respectively), this uniformity does not exist. For the subset of manufacturing industries for which sufficient data are available, the tax substitution with zero CIT shifting results in an aggregate investment increase at 10%. However, individual regions experience increases in investment ranging from 7.5% (West South Central) to 12.5% (New England), even on the intra-industry proportionality assumption. The increase for the South is about equal to that for the nation. However, the South's industrial composition is quite different and its average expansion is explained by the counter-balancing of such regionally important industries as paper and allied products (27% investment increase, of which the South's share is 38%) at the high end, against tobacco (only a 4.8% investment expansion with a Southern share of 67%); chemicals (investment expansion of 3.1%, Southern share of 44%) at the sluggish end of the spectrum. The average expansion of the textile industy, 10.3%, of which the South's share is 65%, also contributes to the South's average standing.

Under the assumption of full forward shifting of the CIT, the national investment expansion in these manufacturing industries is only 1.9%. And in

this case the South's share is somewhat less than average, only 1.7%. The smallest expansion is observed in West South Central (1.5%), the largest again in New England (2.2%). In both cases the New England phenomenon is primarily explained by the rapid expansion of the non-electrical machinery industry (22.6% with zero shifting, 3.4% with unitary shifting of the CIT).

Even granting that the South's share of the investment expansion may be understated by the intra-industry proportionality assumption, it would still appear that this region would not benefit disproportionality from the investment response to the tax substitution, although further analysis might indicate otherwise.

This discussion has examined investment effects by purchaser (investor) industry. While this focus is most important in long run terms, i.e., in terms of differential regional growth, the short run consequences would be most sensitive to increases in activity of investment goods producer industries, the spatial distribution of which will not be identical to that of investor industries. While beyond the scope of this examination, differential regional (and national) consequences of the investment-induced expansion in industrial activity should be treated endogenously and examined explicitly.

That these increases in industrial activity would not be uniformly distributed regionally can be at least suggested by examining the geographic distribution of investment goods producer industries. In Table V - 11 the regional distribution of all industries is presented. Table V - 12 indicates the proportion of investment goods value added contributed by each producer industry. On this basis the regional impacts of the increase in capital goods output can be roughly projected. Ignoring construction, which would probably closely approximate the regional distribution of investment purchases and which accounts for 43% of investment value added, the largest contribution is by the non-electrical machinery industry (25%). On the basis of the spatial distribution of this industry's activity in 1969, 9.0% of this expansion would be concentrated in New England, certainly a disproportionate share.

Again, there is no more justification for the proportionality assumption in allocating producer than user expansion. Certainly, the geographic distribution of users (investors) wil affect the distribution of producers, and even if this were not true, it would be unlikely that marginal output expansions would be distributed spatially in proportion to total output. Different ages of capital stocks and different technologies will serve to alter the geographic distribution of industrial activity in an expansion. Until such processes and factors are considered it will be impossible to adequately predict the full regional consequences of federal policy.

International Trade Effects

Of the debates concerning the potential consequences of a value added tax, none have been more subject to confusion than those concerning the

international trade effects of such a tax. Because the VAT would be rebated on exports and imposed as a border tax on imports, it is often alleged that it would therefore stimulate exports and restrict imports.

These effects are corollaries to the asserted discriminatory consequences of the recent European extensions of the value added tax. Except for Italy, all of the Common Market countries have replaced previously existing indirect business taxes with value added taxes as part of the tax harmonization efforts of the EEC. Similar moves toward value added taxes have taken place in the rest of Europe. These European developments are often cited as a significant contributing cause of the deterioration over the last decade in the U.S. trade position.

The question of the trade consequences of a VAT is simultaneously simple and complex. Most simply stated, there are no trade effects of a VAT per se. Whether changes in export and import patterns will accompany the introduction of a VAT depends upon the relative price consequences of the change in the VAT and in other taxes. If the only consequence of the VAT and simultaneous changes in other taxes, government expenditures and/or deficits is to increase VAT-exclusive prices unchanged, then no trade consequences can be expected (unless domestic disposable income is reduced and the demands for both domestic and foreign products decline). Export prices will be unchanged, and import prices will rise by the same amount as the prices of import-competing goods. More generally, changes in exports will depend upon changes in export prices relative to world market prices, and imports will be affected only if import prices are affected differently than the prices of import-competing goods.

Note that the VAT itself is neutral on these questions. Relative price changes can only result from the other fiscal changes accompanying the introduction of the VAT. In the present case, this means price changes resulting from the reduction and shifting of the corporate profits tax. Only if a) export prices are reduced by the shifting to foreign purchasers of the benefits of the CIT reduction and/or b) import-competing prices rise by less than the VAT rate due to forward CIT shifting will the U.S. trade balance be favorably affected by the tax substitution. Thus, just as in the case of relative price changes generally, the causal burden is on the CIT reduction rather than on the VAT imposition.

If the CIT is not shifted, export prices will be unchanged and U.S. exports will be unaffected. Similarly, import prices will change identically with prices of U.S. import-competing goods. Thus no balance of trade effects would result. Full shifting of the CIT, on the other hand, will reduce prices of U.S. exports relative to world prices and will reduce prices of import-competing goods relative to those of imports. The consequences for the U.S. balance of trade will then depend on the elasticities of export and import demand. As in Table $\overline{V}$ - 13, under a range of reasonable elasticity assumptions the effect would be a reduction in the U.S. trade deficit (increase in surplus) of between $2.5 billion and $4.5 billion.

To place this result in context, even under the most favorable assumption of full CIT shifting, the tax substitution is equivalent to an effective devaluation of only 5 percent. But note that a devaluation would produce the indicated change in trade balance ($2.5 billion to $4.5 billion) regardless of the degree of tax shifting. Thus, devaluation offers a much more certain means of achieving these trade "improvements" (improvements only in the case of prior deficit).

Assuming that the CIT is shifted forward, and hence that the tax substitution serves to stimulate exports (by between $0.6 billion and $2.3 billion) and reduce imports (by between $2.0 billion and $2.2 billion), it is clear that these changes in export and import-competing industrial activity will not be uniformly distributed regionally. This can be suggested by examing Tables $\overline{V}$ - 11 and $\overline{V}$ - 12. It is clear that many of the most important import-competing industries (as measured by imports as a proportion domestic value added) are of predominant importance in the South. Thus, import substitution expansion of textiles and paper products would have pronounced effects in the South. However, on the export side,the only really important Southern industry is chemicals; the South is relatively unrepresented in non- electrical machinery and transportation equipment, other major exporters.

It is impossible to precisely identify the regional effects of the projected trade reactions. First, the trade effects were estimated on the basis of aggregate export and import elasticities and price indices; more disaggregated parameters are simply not available and industry price changes are not identical. And even if these inadequacies were overcome, the problems encountered in regionally allocating increases in capital goods production would also be encountered here. Thus, until this aspect of national-industrial-regional analysis is significantly improved and extended, such discussions can only be suggestive rather than definitive.

Before leaving the issue of international trade effects, one important comment and qualification should be made. It was noted above that the VAT-CIT substitution was at best equivalent to a 5% devaluation. However, these policy changes are equivalent only in terms of the aggregate net change in the balance of trade. But the disaggregated composition of this change in trade balance differs significantly under the two alternatives. In the case of the tax substitution, employing the higher elasticities, exports expand by $2.3 billion and imports contract by $2.2 billion, giving the net improvement in trade balance of $4.5 billion. Under the devaluation, the value of imports is unaffected while exports expand by the full $4.7 billion. The regional consequences would therefore differ significantly between the two cases. Because the stimulus to the South is probably greater through import substitution than through export expansion, this region would be more strongly affected by the tax substitution (assuming shifting of the CIT and a consequent import reduction)

This again raises the general point that examining only aggregate effects

of policies, e.g., changes in net trade balances, does not provide a sufficient basis for evaluating alternative policies, and that this is true at the national as well as regional levels. Without disaggregation both devaluation and a VAT-CIT substitution appear to have identical international trade effects, given certain restrictive assumptions concerning CIT shifting. In fact, this identity is only superficial, and major differences affecting the relative desirability of alternative policies are entirely submerged by this excessive degree of aggregation.

Possible First-Round Wage Adjustments to the VAT-CIT Substitution

Full replacement of the CIT by a consumption-type VAT has been estimated to increase consumption prices by 7.7 percent if the CIT reduction is not shifted forward in lower prices and by 1.7 percent if such shifting does occur. In either case, these price increases imply corresponding reductions in real wages. If the supply of labor is initially assumed to be infinitely elastic at the prevailing real wage, it is possible to estimate the first-round wage adjustment to the tax substitution.

Such wage increases would obviously exert upward pressure on prices. Interindustry variations in this pressure would depend on differences in labor intensity, measured by employee compensation as a proportion of value added (Table V-14). Those industries in which this ratio is highest would be most affected by this wage adjustment. Included in this class would be such important Southern industries as textiles (employee compensation 82 percent of value added) and furniture (84 percent). Contrarily, employee compensation is relatively low in oil and gas (37 percent) and tobacco (33 percent).

Under the assumption of zero CIT shifting, of course, net profits increase by the amount of initial CIT liabilities. It could be assumed that, at least in the first instance, the increase in employee compensation is absorbed by profits. The pressure for further round price changes would then be greater in those industries in which the increase in employee compensation exceeded the original increase in net profit (CIT liability). Original CIT liabilities, the hypothesized labor cost changes, and the difference between the two, as a proportion of value added, are also indicated in Table V-13. Further adjustments would be most severe in industries in which the change in labor cost exceeds the CIT reduction. In no important Southern industry would the net excess in labor cost over CIT be greater than 2 percent. But in such major regional industries as tobacco, chemicals and paper, significant net profit increases would be observed even if the wage increases were fully absorbed by profits.

For manufacturing industries it is possible to estimate the aggregate regional effects of CIT removal and wage increases; as shown in Table V-15. For manufacturing industries as a whole, the original CIT liability greatly exceeds the hypothetical wage adjustment. Nationally, the net increase in profit (CIT liability minus wage adjustment) in manufacturing is 20 percent

of original net capital earnings. This is explained by the predominance of the corporate form in manufacturing, as compared to other sectors.

However, regional variations in this net increase in capital earnings in manufacturing are very great, ranging from only 1.5 percent in New England to 35.3 percent in the West South Central. Significantly, the Southern regions combined register a percentage increase in capital earnings adjusted for wage increases greater than that in any other area, 30.2%. The consequences of this persistent profit increase would be expected to be a greater than average stimulus to investment.

The purpose of this section was to indicate the importance of assessing the differential regional consequences of federal policies which are often not discussed in regional terms or recognized to have significant regional impacts. The identification of these differential regional effects is important for three reasons. First, to the degree to which a particular area is severely impacted by a federal policy change, compensating adjustments in other federal programs are indicated. Secondly, in general there exists more than one federal action which will achieve a desired end, and the choice between these should be made on the basis of differential consequences in other dimensions, of which the regional dimension would not be the least important. And finally, the very effectiveness of a program may itself be affected by its differential regional impacts. For example, a selective tax reduction designed to stimulate demand and employment may be primarily inflationary if, due to the characteristics of those benefited, the expansion is concentrated in industries and areas already relatively fully employed. Only by explicitly assessing the differential effects of alternative policies in disaggregated, including regional, terms is it possible to make rational and effective policy choices.

Conclusion

The idea emphasized in this paper is that the analysis of major federal programs must go beyond superficial description, e.g., the identification of gross revenue sharing awards, and attempt to stipulate the full policy menu and to identify and evaluate the full range of direct and indirect policy consequences. The analysis must be closed: It must deal with the implications of one federal program for another and of federal expenditure and grant programs for federal tax policy. Further, the consequences of any federal policy configuration must be assessed in terms of its implications for the behavior of households, firms and governments. Only when the effects of policy in these highly disaggregated terms can be identified can choices between alternative policies be effectively made.

These consequences of federal policy cannot simply be deduced by working with one policy component or one type of response at a time and then simply adding the findings in some arithmetic manner. Thus, migration behavior has implications for labor markets, labor markets have implications for industrial behavior, and ultimately, the indirect, interactive effects of policy may dominate the direct, superficial consequences which motivated

and justified one federal action over others.

The barrier to this type of analysis is ultimately the lack of a closed, general equilibrium macro-model suited to the evaluation of federal policy in such disaggregated dimensions. The difficulties with aggregative national economic models has certainly not fostered their extension to more disaggregated sectors, e.g., regions. Those attempts which have been made have discovered major barriers in the inapplicability of justified aggregative or national assumptions to disaggregated, subnational analysis. For example, the assumption of fixed import coefficients in national input-output analysis has been found to be almost fatally unjustified at the regional level.

My own guess would be that much greater success will be found in moving from very disaggregated levels of analysis, of the behavior of households and firms, to more aggregated regional and national models, than the reverse. Only when behavior, including that of governments, is explicitly considered and understood, will it be possible to adequately project and evaluate the consequences of public policy.

Table I-1

Regional Federal Income Maintenance Transfers

Region	Federal Contributions Per Capita ($)	Federal Contributions Per Poverty Family ($)	Aggregate Federal Contribution Billion $	(%)*	Allocated Federal Tax Cost Billion $	(%)*	Net Benefit Billion $	(%)*	Income Per Capita ($)
NE	21.36	675	0.243	(5.8%)	0.290	(6.9%)	-0.047	(-1.1%)	3475
MA	18.49	539	0.682	(16.5)	0.991	(23.9)	-0.309	(-7.4)	3567
ENC	13.33	360	0.527	(12.7)	0.936	(22.5)	-0.409	(-9.8)	3354
WNC	18.84	306	0.302	(7.3)	0.260	(6.3)	0.042	(1.0)	3005
SA	15.13	236	0.442	(10.7)	0.466	(11.2)	-0.024	(-0.5)	2665
ESC	23.64	259	0.309	(7.4)	0.139	(3.3)	0.170	(4.1)	2223
WSC	28.56	401	0.549	(13.2)	0.292	(7.0)	0.257	(6.2)	2575
M	19.29	474	0.153	(3.7)	0.122	(3.0)	0.031	(0.7)	2781
P	36.09	1189	0.922	(22.3)	0.634	(15.4)	0.288	(6.9)	3545
US	20.77	430	4.135	(100.0)	4.135	(100.0)	0.000	(0)	3159
South	21.11	293	1.300	(31.3)	0.897	(21.5)	0.403	(9.8)	2544
Non-South	20.62	547	2.835	(68.7)	3.238	(78.5)	-0.403	(-9.8)	3433

*(%) of $4.135 billion.

Table II-1

Regional Non-Welfare Categorical Grant Incidence

Region	Cat. Grants Per Capita ($)	Per Capita Cat. Grants Index* (mean = 1)	Per Capita Normalized Federal Tax Incidence* (mean = 1)	Per Capita Normalized Net Categorical Grant Benefit* (mean = 0)	Actual Net Categorical Grant Benefit Billion $	(%)*
NE	48.68	0.87	1.23	-0.35	-0.224	(-2.0%)
MA	42.56	0.76	1.29	-0.53	-1.089	(-9.8)
ENC	44.27	0.79	1.14	-0.34	-0.759	(-6.8)
WNC	60.27	1.08	0.78	0.30	0.271	(2.4)
SA	50.87	0.91	0.77	0.14	0.236	(2.1)
ESC	65.95	1.18	0.51	0.67	0.490	(4.4)
WSC	51.77	0.93	0.73	0.20	0.213	(1.9)
M	104.03	1.87	0.74	1.13	0.499	(4.5)
P	80.67	1.45	1.19	0.25	0.361	(3.2)
US	55.70	1.00	1.00	0	0	(0)
South	54.35	0.97	0.68	0.29	0.939	(8.4)
Non-South	56.30	1.01	1.14	-0.13	-0.939	(-8.4)

* Grants and tax surcharge normalized at $1 per capita for U.S.

** (%) of $11.088 billion.

Table III-1

Regional General Revenue Sharing Distributions
(Normalized Per Capita, National Mean = 1)

Region	House	Mills	Nixon
NE	0.98	0.85	0.93
MA	1.00	0.89	1.02
ENC	0.97	0.89	0.89
WNC	0.94	1.10	1.06
SA	0.98	1.12	0.95
ESC	1.01	1.38	0.99
WSC	1.04	1.23	0.99
M	1.01	0.91	1.21
P	1.00	0.85	1.12
US	1.00	1.00	1.00
outh	1.00	1.21	.99
Non-South	1.00	0.91	1.01

Table III-2

General Revenue Sharing - Categorical Grant Substitutions, Regional Incidence

	Normalized Per Capita (mean = 0)*			Aggregate Net Benefits, $5 Billion Program					
	HRS-CG	MRS-CG	NRS-CG	HRS-CG		MRS-CG		NRS-CG	
				Billion $	(%)**	Billion $	(%)**	Billion $	(%)**
NE	0.11	-0.03	0.05	0.030	(0.6%)	-0.008	(-0.1%)	0.015	(0.3%)
MA	0.23	0.12	0.26	0.214	(4.2)	0.114	(2.2)	0.240	(4.8)
ENC	0.18	0.10	0.10	0.176	(3.5)	0.096	(1.9)	0.099	(1.9)
WNC	-0.14	0.02	-0.02	-0.056	(-1.1)	0.008	(0.1)	-0.008	(-0.1)
SA	0.06	0.21	0.04	0.047	(0.9)	0.154	(3.0)	0.029	(0.5)
ESC	-0.17	0.19	-0.19	-0.056	(-1.1)	0.064	(1.2)	-0.062	(-1.2)
WSC	0.11	0.30	0.06	0.053	(1.0)	0.144	(2.8)	0.029	(0.5)
M	-0.86	-0.95	-0.66	-0.170	(-3.4)	-0.190	(-3.8)	-0.131	(-2.6)
P	-0.45	-0.60	-0.33	-0.289	(-5.7)	-0.385	(-7.7)	-0.211	(-4.2)
US	0.0	0.0	0.0	0.0	(0)	0.0	(0)	0.0	(0)
South	0.029	0.234	0.031	0.044	(0.8)	0.362	(7.0)	-0.004	(-0.2)
Non-South	-0.013	-0.104	-0.014	-0.044	(-0.8)	-0.362	(-7.0)	0.004	(0.2)

*GRS and categorical grants normalized at $1 per capita for U.S.

**(%) of $5 billion.

Table III-3

Net Regional Transfers of GRS Financed by Income Tax Surcharge

	Normalized Per Capita (mean = 0)*			Aggregate Net Benefits, $5 Billion Program					
	HRS-ST	MRS-ST	NRS-ST	HRS-ST		MRS-ST		NRS-ST	
				Billion $	(%)**	Billion $	(%)**	Billion $	(%)**
NE	-0.24	-0.38	-0.30	-0.070	(-1.4%)	-0.108	(-2.1%)	-0.085	(-1.7%)
MA	-0.30	-0.41	-0.27	-0.275	(-5.5)	-0.375	(-7.5)	-0.250	(-5 0)
ENC	-0.17	-0.25	-0.24	-0.165	(-3.3)	-0.246	(-4.9)	-0.243	(-4.8)
WNC	0.16	0.32	0.28	0.065	(1.3)	0.130	(2.6)	0.113	(2.2)
SA	0.21	0.36	0.19	0.154	(3.0)	0.261	(5.2)	0.136	(2.7)
ESC	0.50	0.87	0.48	0.164	(3.2)	0.285	(5.7)	0.158	(3.1)
WSC	0.31	0.50	0.26	0.149	(2.9)	0.241	(4.8)	0.125	(2.5)
M	0.27	0.18	0.47	0.054	(1.0)	0.035	(0.7)	0.093	(1.8)
P	-0.20	-0.35	-0.08	-0.126	(-2.5)	-0.222	(-4.4)	-0.048	(-0.9)
US	0.0	0.0	0.0	0.0	(0)	0.0	(0)	0.0	(0)
South	0.30	0.51	0.27	0.467	(9.1)	0.787	(15.7)	0.419	(8.3)
Non-South	-0.13	-0.23	-0.12	-0.467	(-9.1)	-0.787	(-15.7)	-0.419	(-8.3)

*GRS and income tax surcharge normalized at $1 per capita for U.S.

**(%) of $5 billion.

Table IV-1

Regional Absolute Net Benefits - in thousand dollars

Special Revenue Sharing Replaces Categorical Grants

	Law	Manpower	Urban	Rural	General Transportation	Urban Transportation	Education	Total
NE	-6.	-2734.	-40575.	2834.	7435.	-73644.	-112.	-106807.
MA	-258.	-28078.	-55048.	-13081.	-6873.	-11946.	-408.	-115692.
ENC	-262.	21158.	-2580.	12507.	-5116.	-5941.	488.	20254.
WNC	-132.	-794.	-24537.	34223.	-3167.	13663.	-181.	19075.
SA	364.	46039.	23161.	-7323.	7021.	37446.	-432.	106276.
ESC	187.	-14834.	13259.	-29967.	-3223.	8043.	-218.	-26753.
WSC	222.	-1274.	48132.	1583.	299.	33010.	-266.	81706.
MT	-66.	-6521.	8661.	6099.	-4217.	11336.	-127.	15165.
PAC	-50.	-12962.	29526.	-6875.	7843.	-11967.	1255.	6770.

Table IV-2

Regional Per Capita Net Benefits - in dollars

Special Revenue Sharing Replaces Categorical Grants

	Law	Manpower	Urban	Rural	General Transportation	Urban Transportation	Education	Total
NE	-0.0005	-0.2365	-3.5554	0.2484	0.6515	-6.4532	-0.0098	-9.3585
MA	-0.0070	-0.7609	-1.4917	-0.3545	-0.1863	-0.3237	-0.0111	-3.1352
ENC	-0.0066	0.5345	-0.0652	0.3160	-0.1293	-0.1501	0.0123	0.5116
WNC	-0.0082	-0.0495	-1.5295	2.1332	-0.1974	0.8517	-0.0113	1.1890
SA	0.0124	1.5730	0.7913	-0.2502	0.2399	1.2794	-0.0147	3.6311
ESC	0.0143	-1.1312	1.0111	-2.2853	-0.2458	0.6134	-0.0166	-2.0401
WSC	0.0115	-0.0662	2.5025	0.0823	0.0155	1.7163	-0.0138	4.2481
MT	-0.0083	-0.8220	1.0917	0.7688	-0.5316	1.4290	-0.0160	1.9116
PAC	-0.0019	-0.5071	1.1552	-0.2690	0.3068	-0.4682	0.0491	0.2649

TABLE V-1

Percentage Reducations in VAT-Exclusive Prices by Aggregated Industry

Industry	S = 25%			S = 50%			S = 75%			S = 100%		
	α=.2	α=.6	α=1.0	α=.2	α=.6	α=1.0	α=.2	α=.6	α=1.0	α=.2	α=.6	α=1.0
1 AGRICULTURE	0.12	0.44	0.94	0.23	0.79	1.51	0.34	1.08	1.91	0.44	1.32	2.20
2 METAL MINING	0.20	0.73	1.52	0.40	1.32	2.51	0.58	1.83	3.22	0.75	2.25	3.75
3 COAL,STN,CLAY MNG+PROD	0.32	1.14	2.36	0.62	2.09	3.94	0.92	2.89	5.08	1.19	3.57	5.96
4 OIL + GAS	0.24	0.83	1.68	0.47	1.55	2.86	0.69	2.16	3.77	0.90	2.71	4.51
5 CONSTRUCTION	0.23	0.83	1.71	0.45	1.51	2.85	0.66	2.09	3.68	0.86	2.59	4.31
6 ORDNANCE	0.28	1.01	2.13	0.54	1.83	3.49	0.79	2.51	4.44	1.03	3.08	5.14
7 FOOD	0.22	0.82	1.75	0.44	1.48	2.83	0.64	2.02	3.58	0.83	2.48	4.14
8 TOBACCO	0.51	1.88	4.11	0.98	3.36	6.52	1.43	4.56	8.11	1.85	5.55	9.24
9 TEXTILES + APPAREL	0.32	1.14	2.40	0.61	2.07	3.94	0.90	2.84	5.01	1.16	3.48	5.80
10 LUMBER,WOOD PRODUCTS	0.27	0.93	1.86	0.53	1.73	3.20	0.77	2.42	4.23	1.01	3.03	5.05
11 FURNITURE + FIXTURES	0.31	1.12	2.34	0.61	2.03	3.87	0.88	2.80	4.94	1.15	3.45	5.74
12 PAPER + PRODUCTS	0.37	1.39	2.88	0.76	2.53	4.78	1.10	3.49	6.14	1.43	4.29	7.16
13 PRINTING + PUBLISHING	0.31	1.50	3.18	0.80	2.72	5.19	1.17	3.72	6.59	1.52	4.57	7.61
14 CHEM.,PLAST.,DRUGS,PNT	0.58	2.10	4.49	1.12	3.78	7.27	1.62	5.17	9.17	2.11	6.33	10.54
15 RUBBER + LEATHER	0.35	1.27	2.69	0.68	2.30	4.39	0.99	3.14	5.56	1.29	3.85	6.42
16 FOOTWEAR	0.26	0.95	2.00	0.51	1.73	3.29	0.75	2.37	4.19	0.97	2.92	4.86
17 PRIMARY METAL	0.26	0.93	1.92	0.51	1.70	3.20	0.74	2.34	4.13	0.97	2.90	4.84
18 FABRICATED METAL	0.35	1.27	2.67	0.68	2.29	4.37	1.00	3.15	5.57	1.29	3.87	6.45
19 NONELECT. MACHINERY	0.34	1.23	2.60	0.66	2.22	4.24	0.96	3.04	5.37	1.24	3.73	6.23
20 ELECTRICAL EQUIPMENT	0.38	1.36	2.90	0.73	2.46	4.71	1.06	3.36	5.96	1.37	4.12	6.87
21 TRANSP. EQUIPMENT	0.33	1.20	2.54	0.65	2.18	4.15	0.94	2.99	5.27	1.22	3.66	6.10
22 INSTRUMENTS	0.47	1.73	3.68	0.93	3.13	5.99	1.35	4.28	7.58	1.75	5.24	8.74
23 MISC. MANUFACTURING	0.29	1.03	2.16	0.56	1.87	3.55	0.81	2.57	4.54	1.05	3.15	5.27
24 TRANSP. + WAREHOUSING	0.18	0.71	1.74	0.35	1.24	2.49	0.51	1.65	2.96	0.66	1.98	3.30
25 COMMUNICATIONS	0.64	2.34	4.96	1.25	4.24	8.10	1.83	5.80	10.27	2.37	7.12	11.86
26 UTILITIES	0.51	1.83	3.73	1.00	3.35	6.28	1.47	4.63	8.14	1.91	5.74	9.56
27 FINANCE + INSURANCE	0.89	3.32	7.26	1.73	5.93	11.52	2.52	8.04	14.32	3.26	9.75	16.30
28 REAL ESTATE + RENTAL	0.12	0.43	0.91	0.23	0.78	1.49	0.34	1.07	1.89	0.44	1.31	2.19
29 HOTELS + SERVICES	0.24	0.85	1.78	0.46	1.54	2.93	0.67	2.12	3.74	0.87	2.61	4.34
30 AUTO REPAIR + SERVICES	0.20	0.73	1.54	0.39	1.33	2.53	0.57	1.82	3.21	0.74	2.23	3.72
31 AMUSEMENTS	0.16	0.59	1.44	0.30	1.04	2.06	0.44	1.40	2.50	0.56	1.69	2.82
32 MED.,ED.SERV.+NONPROF.	0.12	0.44	0.93	0.23	0.79	1.52	0.34	1.09	1.92	0.45	1.34	2.23
33 MISCELLANEOUS	0.06	0.21	0.45	0.11	0.39	0.73	0.17	0.53	0.93	0.22	0.65	1.08
34 IMPORTS	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

S : Degree of CIT reduction. α : CIT shifting parameter.

1. Reduction in prices are for all components of final demand in VAT-exclusive producer prices.
2. If α = 0, there are no VAT-exclusive producer price changes.
3. Industries correspend closely to the SIC two-digit classification and to that employed by Aaron.

TABLE V-2

VAT Rates and Revenues Under Alternative CIT Reduction and Shifting Assumptions

(% or billions dollars)

Degree of CIT Reduction

CIT Shifting Parameter α = \ S =	25% Rate (Revenue)	50% Rate (Revenue)	75% Rate (Revenue)	85% Rate (Revenue)	90% Rate (Revenue)	100% Rate (Revenue)	Range in %
.0	1.92 (10.67)	3.84 (21.34)	5.76 (32.01)	6.53 (36.28)	6.91 (38.41)	7.68 (42.68)	300 (300)
.2	2.04 (11.33)	3.99 (22.02)	5.83 (32.14)	6.55 (36.04)	6.90 (37.95)	7.59 (41.72)	271 (268)
.4	2.19 (12.11)	4.15 (22.79)	5.91 (32.28)	6.57 (35.79)	6.89 (37.48)	7.50 (40.77)	242 (237)
.6	2.38 (13.07)	4.34 (23.66)	6.00 (32.43)	6.59 (35.53)	6.87 (37.00)	7.41 (39.81)	212 (205)
.8	2.61 (14.27)	4.56 (24.65)	6.09 (32.59)	6.61 (35.26)	6.86 (36.52)	7.32 (38.86)	175 (172)
1.0	2.91 (15.83)	4.82 (25.80)	6.19 (32.76)	6.63 (34.99)	6.84 (36.02)	7.22 (37.91)	148 (139)
Range in %	51.6 (48.4)	25.5 (20.9)	7.5 (2.3)	1.5 (-3.6)	-1.0 (-6.2)	-5.9 (-11.2)	

α = CIT shifting parameter

S = Reduction in corporate income tax

Figure in parentheses = VAT revenue in billions dollars

VAT Revenue = Net loss in government revenue before the VAT introduction, but after reduction of the CIT.

TABLE V-3

Alternative Aggregate Price Indexes for Final Demands

CIT. Reduction	CIT Shifting	VAT Rate	Price Indices					
S(%) =	α =	z(%) =	P_C(%) =	P_I(%) =	P_R(%) =	P_E(%) =	P_F(%) =	P_{SL}(%) =
25%	.0	1.92	101.92	100.00	100.00	100.00	100.00	100.00
	.2	2.04	101.76	99.71	99.76	99.71	99.71	99.74
	.4	2.19	101.57	99.36	99.48	99.36	99.37	99.43
	.6	2.38	101.34	98.93	99.15	98.94	98.96	99.06
	.8	2.61	101.05	98.42	98.75	98.42	98.44	98.61
	1.0	2.91	100.69	97.76	98.24	97.74	97.79	98.04
50%	.0	3.84	103.84	100.00	100.00	100.00	100.00	100.00
	.2	3.99	103.42	99.43	99.54	99.44	99.44	99.49
	.4	4.15	102.95	98.79	99.02	98.80	98.81	98.92
	.6	4.34	102.42	98.07	98.45	98.09	98.11	98.28
	.8	4.56	101.82	97.25	97.81	97.28	97.31	97.57
	1.0	4.82	101.14	96.32	97.08	96.34	96.39	96.75
75%	.0	5.76	105.76	100.00	100.00	100.00	100.00	100.00
	.2	5.83	104.09	99.16	99.32	99.18	99.18	99.25
	.4	5.91	104.18	98.28	98.61	98.30	98.32	98.46
	.6	6.00	103.32	97.34	97.86	97.38	97.41	97.63
	.8	6.09	102.41	96.36	97.07	96.41	96.44	96.75
	1.0	6.19	101.45	95.31	96.23	95.37	95.42	95.83
100%	.0	7.68	107.68	100.00	100.00	100.00	100.00	100.00
	.2	7.59	106.48	98.91	99.11	98.93	98.94	99.03
	.4	7.50	105.28	97.82	98.24	97.86	97.88	98.05
	.6	7.41	104.08	96.73	97.36	96.79	96.82	97.08
	.8	7.32	102.88	95.64	96.47	95.72	95.75	96.10
	1.0	7.22	101.68	94.55	95.59	94.65	94.69	95.13

Note: S = reduction in corporate income tax; α = shifting parameter; z = VAT rate; P = aggregate price index; C = personal consumption expenditures; I = private nonresidential fixed investment; R = private residential structures; E = exports; F = federal government expenditures; SL = expenditures by state and local governments.

TABLE V-4

Dual VAT Rates : Incidence

Zero CIT Shifting ($\alpha = 0$)

Income Class	Profit Increase $\Delta P/Y$	Consumption Expenditure Increase: $(\Delta C/C)$ and $\Delta C/Y$			Net Incidence $\Delta C/Y - \Delta P/Y$		
		Full Rate	Half Rate	Zero Rate	Full Rate	Half Rate	Zero Rate
$ - 3	0.041	(0.0768) 0.091	(0.0697) 0.082	(0.0543) 0.064	0.050	0.041	0.023
3 - 5	0.054	(0.0768) 0.075	(0.0742) 0.072	(0.0687) 0.067	0.021	0.018	0.013
5 - 7.5	0.014	(0.0768) 0.074	(0.0755) 0.072	(0.0727) 0.069	0.060	0.058	0.055
7.5 - 10	0.052	(0.0768) 0.069	(0.0768) 0.069	(0.0768) 0.069	0.017	0.017	0.017
10 - 15	0.088	(0.0768) 0.065	(0.0782) 0.066	(0.0813) 0.069	-0.023	-0.022	-0.019
15 -	0.124	(0.0768) 0.053	(0.0798) 0.055	(0.0861) 0.059	-0.071	-0.069	-0.065

	VAT Rates		Gini Coefficients
	Non-Exempt	Exempt	g
Full Rate :	7.68%	7.68%	0.397
Half Rate :	10.48	5.24	0.395
Zero Rate :	16.48	0.00	0.394
Pre-VAT :			0.374

Consumption Price Index (Independent of Exemption) : $P_C = 1.0768$.

Unitary Relative Price Elasticity Assumed (1960-61 to 1969 expenditure adjustment).

TABLE V-5

Dual VAT Rates: Incidence

Full CIT Shifting ($\alpha = 1$)

Income Class	Investment Cost Reduction $\Delta I/Y$	Consumption Expenditure Increase : $(\Delta C/C)$ and $\Delta C/Y$			Net Incidence $\Delta C/Y - \Delta I/Y$		
		Full Rate	Half Rate	Zero Rate	Full Rate	Half Rate	Zero Rate
3 - 3	0.0065	(0.023)9.027	(0.018)0.021	(0.004)0.004	0.021	0.014	-0.002
3 - 5	0.0086	(0.021)0.021	(0.020)0.020	(0.015)0.015	0.012	0.011	0.006
5 - 7.5	0.0022	(0.021)0.020	(0.021)0.020	(0.019)0.018	0.018	0.018	0.016
7.5 - 10	0.0083	(0.021)0.019	(0.022)0.020	(0.023)0.021	0.011	0.012	0.012
10 - 15	0.0140	(0.022)0.019	(0.024)0.021	(0.028)0.023	0.005	0.007	0.009
15 -	0.0198	(0.025)0.017	(0.028)0.019	(0.034)0.023	-0.003	-0.001	0.004

	VAT Rates		Gini Coefficients
	Non-Exempt	Exempt	g
Full Rate :	7.22%	7.22%	0.380
Half Rate :	9.89	4.95	0.377
Zero Rate :	15.69	0	0.377
Pre-VAT :			0.374

Consumption Price Index (Independent of Exemption) : $P_C = 1.0168$.

Unitary Relative Price Elasticity Assumed (1960-61 to 1969 expenditure adjustment).

TABLE V-6

Changes in Regional Tax Liabilities: VAT-CIT Substitution
(1969)

Region	Retail Sales (%)	Dividends Received (%)	S=100%, α= 0, β= 1				S=100%, α= 1, β= 1				ΔC-ΔCIT / Y	ΔC-ΔI / Y	Disposable Income Per Capita
			ΔC bil.	ΔCIT bil.	ΔC-ΔCIT bil.	(ΔC-ΔCIT)/N dollars	ΔC bil.	ΔI bil.	ΔC-ΔI bil.	(ΔC-ΔI)/N dollars			
NE	6.2	10.7	2.62	4.58	-1.95	-167	0.57	.73	-.16	-13	-0.049	-0.004	3370
MA	18.7	28.2	7.97	12.04	-4.07	-110	1.74	1.92	-.18	-5	-0.032	-0.001	3490
ENC	20.6	18.8	8.78	8.04	0.75	19	1.92	1.28	0.64	16	0.006	0.005	3330
WNC	8.4	5.8	3.59	2.45	1.14	70	.79	.39	0.39	24	0.023	0.008	3020
SA	14.4	13.0	6.15	5.57	0.58	19	1.34	.89	0.46	15	0.007	0.005	2760
ESC	5.2	2.8	2.21	1.20	1.01	79	.48	.19	0.29	23	0.034	0.010	2340
WSC	8.7	5.2	3.70	2.24	1.46	76	.81	.36	0.45	24	0.028	0.009	2710
M	4.0	2.8	1.69	1.19	0.50	62	.37	.19	0.18	22	0.022	0.008	2830
P	14.0	12.6	5.96	5.39	0.57	22	1.30	.86	0.44	17	0.006	0.005	3400
US	100.0	100.0	42.68	42.68	0.00	00	9.34	6.82	2.52	13	0.0	0.004	3120
South	28.3	21.0	12.06	9.01	3.05	49	2.63	1.44	1.20	19	0.021	0.007	2660
Non-S	71.7	79.0	30.62	33.67	-3.05	-22	6.71	5.38	1.32	11	-0.007	0.003	3320

1. S = CIT reduction, α = CIT shifting, and β = VAT shifting.

2. ΔC = increase in nominal consumption expenditures, ΔCIT = CIT reduction, N = population, ΔI = savings in investment expenditures, and Y = disposable income.

3. Data on 1969 retail sales are taken from "Monthly Retail Trade," January 1970, Bureau of the Census, Dept. of Commerce

4. Data on dividends received are taken from 1969 Individual Income Tax Returns, Internal Revenue Service, Dept. of Treasury.

5. Data on population are taken from 1971 Statistical Abstract of the United States, B. of Census, Dept. of Commerce.

6. Figures for the United States are computed from 1969 U.S. input-output model.

TABLE V-7

Investment Effects of the Tax Substitution, Cash-Flow Approach:
Zero CIT Shifting

(millions of dollars or %)

Industry	CIT Liability ($)	Dividend Increase ($)	Increase in Net Cash Flow ($)	Cumulative Increase in Gross Investment ($)	Short-run Investment Effect (%)
Food and Tobacco	1727	592	1135	212	4.8
Textiles and Apparel	964	282	682	483	10.3
Paper & Allied Products	873	250	623	1450	26.9
Chemicals	3182	954	2228	1226	5.6
Petroleum and Coal	692	207	485	638	3.1
Rubber	454	64	390	293	11.4
Stone, Clay and Glass	630	190	440	543	9.2
Iron and Steel	818	314	504	496	2.2
Nonferrous Metals	257	99	158	530	7.3
Machinery, except Electrical	2446	939	1507	2263	22.6
Electrical Machinery	2182	838	1344	2753	8.6
Manufacturing	21923	6351	15572	20580	11.5
All Industries	42680	18429	24251	32050	5.3

Note: (1) Estimation of dividend effect is based on J.A. Brittain, Corporate Dividend Policy, Table 27 and Equation (4-10); (2) Estimation of investment is based on Table VII-5 and Equation (1) in Table VII-2 of J.R. Meyer and R.R. Glauber, Investment Decisions, Economic Forecasting, and Public Policy; (3) Percent increase in gross investment is computed by means of (short-run) cash-flow elasticity indicated in Table VII-6 of Meyer and Glauber; (4) Cash flow = net profits - dividend + depreciation.

TABLE V-8

Investment Effects of the Tax Substitution, Cash-Flow Approach:
Full CIT Shifting

(millions of dollars or %)

Industry	Corporate Cash Flow ($)	Reduction in Price of Investment Goods (%)	Increase in Real Net Corporate Cash Flow ($)	Increase in Corporate Gross Invest. ($)	Increase in Total Gross Investment ($)	Short-run Investment Effect (%)
Food and Tobacco	3819	5.95	242	45	57	1.3
Textiles and Apparel	1614	5.42	93	66	68	1.5
Paper & Allied Products	1917	5.32	108	251	251	4.7
Chemicals	4614	5.81	284	156	160	.7
Petroleum and Coal	395	5.26	220	289	307	1.5
Rubber	946	5.72	57	43	43	1.7
Stone, Clay and Glass	1416	5.78	87	107	112	1.9
Iron and Steel	2903	5.75	177	174	175	.8
Nonferrous Metals	915	5.60	54	182	182	2.5
Machinery, except Elec'l.	3792	5.52	221	332	343	3.4
Electrical Machinery	2845	5.71	172	353	356	1.1
Manufacturing	33995	5.50	1979	2615	2672	1.5
All Industries	73627	5.45	4244	5609	8980	1.5

Note: (1) See Table 9-1, note (2), for the estimation of corporate gross investment; (2) Increase in total gross investment (corporate and non-corporate) is approximated by multiplying increase in corporate gross investment by one plus the ratio between depreciation in the two sectors.

TABLE V-9

Investment Expansion by Region
Zero CIT Shifting
(in million $)

	NE	MA	ENC	WNC	SA	ESC	WSC	MT	PAC	US	South	Non-South
Food	4.52	13.83	19.96	11.89	13.10	4.11	7.23	3.94	13.35	91.93	24.44	67.49
Tobacco	---	.22	.06	---	2.14	.46	---	---	---	2.36	2.60	.28
Textile Prod.	6.68	9.72	2.02	.31	55.87	9.26	1.12	---	---	84.98	66.25	18.73
Apparel	.83	16.03	3.34	1.47	4.61	2.25	1.39	.23	1.85	30.15	8.25	23.75
Paper	30.42	60.26	80.40	21.22	75.27	37.88	32.33	3.07	40.78	380.63	144.48	236.15
Chemical	3.76	25.43	26.59	5.06	30.06	12.76	48.29	1.66	5.51	159.22	91.11	68.11
Petrol. and Coal	.23	2.90	8.34	1.26	.54	1.34	12.65	.46	5.53	33.25	14.53	18.72
Rubber and Plastic	8.63	13.52	35.11	5.22	9.42	14.14	5.57	---	8.63	100.24	29.13	71.11
Stone, Clay, and Glass	2.91	18.15	22.75	4.49	11.70	6.52	4.85	3.26	8.84	83.47	23.07	60.40
Primary Metal	3.62	37.94	83.01	3.28	12.61	8.47	7.83	6.31	8.71	172.08	28.91	143.17
Fabric. Metal	4.86	13.55	34.37	3.98	5.84	4.62	3.84	.60	6.82	78.48	14.30	64.18
Machinery	33.92	91.76	165.14	30.40	25.85	12.63	20.25	4.84	36.36	421.15	58.73	362.42
Electrical Equipment	14.27	32.14	38.47	6.14	13.51	7.18	10.31	2.77	16.33	141.12	31.00	110.12
Total	114.65	335.45	519.56	94.72	260.52	121.62	155.66	27.14	152.81	1781.43	536.80	1244.63
Percentage Change for Total	12.5	10.3	10.1	10.6	10.3	10.4	7.5	7.9	10.2	10.0	9.3	10.3

Figures are obtained by applying percentage changes in gross investment by industry, estimated from the 1969 input-output tables, to 1969 capital expenditures by industry over regions, taken from 1969 Annual Survey of Manufacturing, Census of Bureau, Department of Commerce.

TABLE V-10

Investment Expansion by Region
Unitary CIT Shifting
(in million $)

	NE	MA	ENC	WNC	SA	ESC	WSC	MT	PAC	US	South	Non-South
Food	1.22	3.75	5.41	3.22	3.55	1.11	1.96	1.07	3.62	24.91	6.62	18.29
Tobacco	---	.06	.02	---	.58	.12	---	---	---	.78	.70	.08
Textile Products	.97	1.42	.29	.04	8.14	1.35	.16	---	---	12.37	9.65	2.72
Apparel	.12	2.33	.49	.21	.67	.33	.20	.03	.27	4.65	1.20	3.45
Paper	5.32	10.53	14.05	3.71	13.15	6.62	5.65	.54	7.13	66.70	25.42	41.28
Chemical	.47	3.18	3.32	.63	3.76	1.59	6.04	.21	.70	19.90	11.39	8.51
Petrol. and Coal	.11	1.40	4.04	.61	.26	.65	6.12	.22	2.68	16.09	7.03	9.06
Rubber and Plastic	1.29	2.02	5.24	.78	1.40	2.11	.83	---	1.29	14.96	4.34	10.62
Stone, Clay, and Glass	.60	3.75	4.70	.93	2.42	1.35	1.00	.67	1.83	17.25	4.77	12.48
Primary Metal	1.25	13.06	28.58	1.13	4.34	2.92	2.70	2.17	3.00	59.15	9.96	49.19
Fabric. Metal	1.67	4.66	11.83	1.37	2.01	1.59	1.32	.21	2.35	27.01	4.92	22.09
Machinery	5.10	13.80	24.84	4.57	3.89	1.90	3.05	.73	5.47	63.35	8.84	54.51
Electrical Equip.	1.82	4.11	4.92	.79	1.73	.92	1.32	.35	2.09	18.05	3.97	14.08
Total	19.94	64.07	107.73	17.99	45.90	22.56	30.35	6.20	30.43	345.17	98.81	246.36
Percentage Change for Total	2.2	2.0	2.1	2.0	1.8	1.9	1.5	1.8	2.0	1.9	1.7	2.0

Table V-11

Value Added in Selected Industries, 1969, By Region

Percentage Distribution

	NE	MA	ENC	WNC	SA	ESC	WSC	MT	PAC	SUM
Ordnance	6.36	4.62	11.92	12.74	7.73	3.01	5.42	5.16	42.18	99.14
Food	3.64	17.34	24.26	11.78	11.32	5.61	7.84	2.99	14.60	99.38
Tobacco	.20	5.39	1.24	---	67.54	*	---	---	---	74.37
Textiles	9.13	16.73	3.45	.47	58.23	8.10	1.38	---	---	97.49
Apparel	5.39	43.61	8.92	3.74	16.03	9.31	5.22	.70	6.49	99.41
Lumber and Wood	4.27	6.07	10.86	3.58	13.38	9.49	8.30	5.99	37.42	99.36
Furniture	5.04	16.82	24.05	4.39	23.55	8.13	5.28	.87	11.36	99.49
Paper	10.13	18.34	23.33	5.38	15.98	7.70	7.08	.94	11.22	100.10
Printing	6.50	33.35	24.76	7.97	8.86	3.34	4.53	2.04	9.74	101.07
Plastics and Chemicals	4.00	25.67	20.79	5.21	16.08	9.15	12.93	1.07	6.10	101.00
Petroleum	.76	11.39	17.66	6.30	3.01	2.50	40.60	2.68	14.73	99.63
Rubber	10.44	16.82	37.76	6.46	8.31	6.93	4.51	---	*	91.23
Leather	26.60	23.13	15.39	*	6.52	*	*	---	*	71.64
Stone, Clay and Glass	4.62	21.44	26.18	7.22	13.77	5.39	7.28	2.60	10.77	99.27
Primary Metal	4.38	25.23	40.22	2.54	6.53	5.95	4.31	3.90	7.01	100.07
Fabric. Metal	7.31	19.01	41.08	5.49	6.33	4.33	5.20	1.02	9.76	99.53
Machinery	8.74	18.48	42.12	8.97	4.39	3.51	4.41	1.63	7.83	100.08
Electrical Equip.	9.34	24.34	30.81	4.93	7.71	5.10	3.31	1.72	12.74	100.00
Transportation	5.50	11.85	39.53	6.99	6.55	2.46	6.17	6.36	17.42	102.83
Instrument	14.80	45.19	17.61	4.61	3.75	1.46	2.72	1.77	7.31	99.22
Misc. Manufacturing	17.93	32.46	19.85	5.39	6.28	3.98	3.12	0.90	9.80	99.71

*Withheld to avoid disclosing individual firms.

Table V-12

Industry Share of Investment and Export Value Added and Imports as

a Proportion of Domestic Value Added

	Share of Plant & Equipment Value Added	Share of Export Value Added	Imports as Percentage of Value Added
1. Agriculture		9.11	6.81
2. Metal Mining		.23	79.75
3. Coal, Stn, Clay Mng & Prod		1.98	6.15
4. Oil & Gas		.52	26.73
5. Construction	42.99		
6. Ordnance		3.07	11.75
7. Food		2.74	14.15
8. Tobacco		.40	.63
9. Textiles & Apparel	.07	.98	15.96
10. Lumber, Wood Products	.02	1.05	19.68
11. Furniture & Fixtures		.09	5.22
12. Paper & Products		1.43	15.99
13. Printing & Publishing		.66	1.34
14. Chem., Plast.,Drugs, Pnt		7.82	5.32
15. Rubber & Leather	.18	.80	7.68
16. Footwear	.02	.04	24.83
17. Primary Metal		3.36	14.12
18. Fabricated Metal	1.75	2.51	2.34
19. Nonelect. Machinery	24.87	11.36	6.86
20. Electrical Equipment	6.84	5.99	8.21
21. Transp. Equipment	12.57	12.29	19.72
22. Instruments	3.49	2.22	8.22
23. Misc. Manufacturing	.82	.69	23.57
24. Transp. & Warehousing	1.51	13.31	7.75
25. Communications	1.79	.41	
26. Utilities		.02	.29
27. Finance & Insurance			1.40
28. Real Estate & Rental		4.47	
29. Hotels & Services		11.10	
30. Auto Repair & Services			
31. Amusements		.08	
32. Med., Ed. Serv. & Nonprof.			
33. Wholesale & Retail		1.12	

TABLE V-13 Hypothetical Changes in U.S. Trade Balance

(billions dollars)

Degree of CIT Reduction	CIT Shifting Parameter	Export Price Change	Import-Competing Price Change	Change in Exports	Change in Imports	Change in Trade Balance	Change in Exports	Change in Imports	Change in Trade Balance
S(%) =	α =	(%) $\frac{\Delta P_x}{P_x}$	(%) $\frac{\Delta P_d}{P_d}$	(ϵ_x = -1.24)	(ϵ_m = -.88)		(ϵ_x = -2.00)	(ϵ_m = -1.00)	
				ΔV_x	$\Delta V_{m'}$ ΔV_m	ΔV	ΔV_x	$\Delta V_{m'}$ ΔV_m	ΔV
	.2	.29	.31	.030	.007 .113	.136	.126	.011 .128	.243
25	.6	1.06	1.14	.111	.025 .412	.498	.461	.041 .707	1.127
	1.0	2.26	2.39	.236	.053 .879	1.298	.983	.086 .999	1.896
	.2	.56	.61	.058	.013 .218	.266	.244	.022 .248	.470
50	.6	1.91	2.06	.199	.045 .743	.897	.831	.073 .844	1.602
	1.0	3.66	3.92	.382	.085 1.424	1.721	1.592	.137 1.618	3.073
	.2	.82	.89	.086	.020 .319	.385	.357	.032 .362	.687
75	.6	2.62	2.83	.274	.062 1.019	1.231	1.140	.099 1.158	2.199
	1.0	4.63	5.00	.483	.107 1.801	2.177	2.014	.171 2.046	3.889
	.2	1.07	1.16	.112	.026 .416	.502	.465	.041 .473	.897
100	.6	3.21	3.48	.335	.075 1.249	1.509	1.396	.121 1.419	2.694
	1.0	5.35	5.81	.559	.122 2.081	2.518	2.327	.196 2.365	4.496

Note : (1) $\Delta V_x = V_x \frac{\Delta P_x}{P_x}(1+\epsilon_x)$, where V_x = exports of goods and services exclusive of investment income, ΔP_x = change in export price, and ϵ_x = relative price elasticity of exports.

(2) $\Delta V_m = -V_m \frac{\Delta P_d}{P_d}\epsilon_m$, where V_m = imports of goods and services exclusive of investment income, ΔP_d = change in import price, and ϵ_m = relative price elasticity of imports.

(3) $\Delta V_{m'} = \alpha(V_x \frac{\Delta P_x}{P_x}\epsilon_x)(1 - \frac{\Delta P_d}{P_d}\epsilon_m)$ = Induced imports due to exports, where α = total import coefficient per unit of exports = .0447.

(4) $\Delta V = \Delta V_x + \Delta V_m - \Delta V_{m'}$.

(5) Changes in aggregated price indexes ΔP_x and ΔP_d are obtained by weighting changes in pre-VAT producer price by 1969 exports and imports, respectively.

(6) V_x = 43.5 billions and V_m = 44.2 billions.

TABLE V-14

Tax Liability, Labor Compensation and Labor Cost-CIT Differentials
Zero CIT Shifting

	VAT-CIT Difference (Billions of $)	VAT/CIT Ratio	Labor Compensation (Ratio to Value Added)	CIT (Ratio to Value Added)	Labor Cost Increase (Ratio to Value Added)	Labor Cost Increase - CIT (Ratio to Value Added)
1 AGRICULTURE	1.5963	13.4	0.1321	0.0042	0.0101	0.0059
2 METAL MINING	-0.0138	0.6	0.8368	0.0343	0.0643	0.0300
3 COAL,STN,CLAY MNG&PROD	-0.3521	0.5	0.7205	0.0612	0.0553	-0.0057
4 OIL & GAS	-0.0153	1.0	0.3726	0.0619	0.0286	-0.0331
5 CONSTRUCTION	0.0313	1.0	0.7290	0.0236	0.0560	0.0324
6 ORDNANCE	-0.1180	0.2	0.8650	0.0429	0.0664	0.0235
7 FOOD	0.4450	1.4	0.7459	0.0555	0.0573	0.0018
8 TOBACCO	-0.2433	0.5	0.3833	0.3054	0.0294	-0.2758
9 TEXTILES & APPAREL	0.0984	1.1	0.8207	0.0589	0.0630	0.0041
10 LUMBER,WOOD PRODUCTS	-0.2084	0.5	0.6369	0.0605	0.0489	-0.0115
11 FURNITURE & FIXTURES	-0.0705	0.7	0.8385	0.0678	0.0644	-0.0033
12 PAPER & PRODUCTS	-0.4190	0.5	0.6904	0.0873	0.0530	-0.0342
13 PRINTING & PUBLISHING	-0.4139	0.6	0.7630	0.0875	0.0586	-0.0288
14 CHEM.,PLAST.,DRUGS,PNT	-2.1641	0.3	0.5984	0.1500	0.0460	-0.1039
15 RUBBER & LEATHER	-0.0714	0.8	0.7925	0.0649	0.0609	-0.0039
16 FOOTWEAR	0.0338	1.3	0.8500	0.0515	0.0653	0.0137
17 PRIMARY METAL	-0.4957	0.5	0.7558	0.0527	0.0580	0.0053
18 FABRICATED METAL	-0.7531	0.4	0.4536	0.0485	0.0348	-0.0135
19 NONELECT. MACHINERY	-1.7569	0.3	0.6781	0.0855	0.0521	-0.0333
20 ELECTRICAL EQUIPMENT	-1.3214	0.4	0.7818	0.0876	0.0600	-0.0275
21 TRANSP. EQUIPMENT	-1.5376	0.4	0.7672	0.0908	0.0589	-0.0318
22 INSTRUMENTS	-0.6191	0.3	0.6511	0.1404	0.0500	-0.0903
23 MISC. MANUFACTURING	0.0029	1.0	0.8503	0.0577	0.0653	0.0076
24 TRANSP. & WAREHOUSING	0.7282	2.0	0.7318	0.0230	0.0562	0.0332
25 COMMUNICATIONS	-1.5929	0.4	0.5049	0.1465	0.0388	-0.1076
26 UTILITIES	-1.4341	0.4	0.3270	0.1335	0.0251	-0.1083
27 FINANCE & INSURANCE	-5.1543	0.3	0.8211	0.2507	0.0631	-0.1876
28 REAL ESTATE & RENTAL	4.0801	6.2	0.0278	0.0110	0.0021	-0.0088
29 HOTELS & SERVICES	1.3292	2.8	0.6034	0.0187	0.0463	0.0276
30 AUTO REPAIR & SERVICES	0.4182	3.3	0.4846	0.0178	0.0372	0.0194
31 AMUSEMENTS	0.2744	9.4	0.7100	0.0064	0.0545	0.0481
32 MED.,ED.SERV.&NONPROF.	2.5117		0.6638	-0.0011	0.0510	0.0522
33 WHOLESALE & RETAIL	3.2107	1.6	0.7313	0.0477	0.0562	0.0084

Table V-15

Regional Labor Cost-CIT Differentials for Manufacturing
(billions)

	Net Value-Added (1)	Employee Compensation (2)	Net Capital Income (3)	CIT Reduction (4)	Increase in Employee Compensation (5)	Increase in Profit (6)=(4)-(5)	Increase in Profit Relative to Net Capital Income in % (7)=(6)/(3)
U.S.	221.3	179.9	41.4	21.9	13.8	8.1	19.6
N.E.	15.3	13.3	2.0	1.05	1.02	.03	1.5
M.A.	47.2	39.8	7.4	3.93	3.05	.88	11.9
E.N.C.	63.3	52.6	10.7	5.67	4.03	1.64	15.3
W.N.C.	14.3	11.0	3.3	1.77	.84	.93	27.8
S.A.	25.1	19.8	5.3	2.79	1.52	1.27	24.1
E.S.C.	12.1	8.7	3.4	1.80	.67	1.13	33.2
W.S.C.	14.4	10.0	4.3	2.30	.77	1.53	35.3
MT	4.1	3.1	1.0	.52	.24	.28	28.3
PAC	25.6	21.8	3.9	2.05	1.67	.38	9.8

(1) Full CIT reduction, zero CIT shifting, and full VAT shifting are assumed.

(2) Regional Allocation of Net Value Added is based on value-added data from 1969 Annual Survey of Manufactures, Bureau of the Census, Department of Commerce.

(3) Regional Allocation of Employee Compensation is based on data on 1969 Personal Income by Region and on Income and Employment by Industry, Survey of Current Business, July 1970, Office of Business Economics, Dept. of Commerce.

(4) Net Capital Income = Net Value Added - Employee Compensation.

(5) CIT Reduction is allocated by the Distribution of Net Capital Income.

(6) Employee Compensation is assumed to increase by the VAT rate, 7.68%.

POLICY IMPLICATIONS OF INTERMETROPOLITAN MIGRATION FLOWS

WILLIAM ALONSO

In the foreseeable future it appears that the greatest force molding the overall population distribution of the United States will not be, as is commonly supposed, the migration out of rural areas and small cities into metropolitan areas. Rather, the dominant force will be the migratory cross-currents among metropolitan areas. These areas, which already hold over two-thirds of the nation's population, grew by 16.5% in the 1960's, compared to 6.8% for the rest of the country. But their growth was largely (73%) from natural increase, while immigrants from non-metropolitan areas accounted for less than 12% of their increase. By contrast, 77% of net civilian immigration into this country ended in the metropolitan areas, accounting for 15% of their total growth -- a substantially greater portion of total increase than that of domestic migration. Thus metropolitan areas, which already dominate the national pattern of population distribution, grew primarily from their own natural increase and from direct immigration from abroad.

What is of interest here is how the population distributes itself among the set of metropolitan areas. Their growth varied widely in the 60's, from over 100% to a net loss. This was the result of great variations in their rates of natural increase and of migration. Natural increase in the 1960's varied astonishingly among metropolitan areas by a factor of 23, from 1.3% in Scranton, Pa., to 29.1% in Laredo, Texas. Similarly, 39% of them experienced a net outward migration in the 1960's, and even four of the largest ten were net losers in migratory exchanges.

In the past the relative variation in the growth rates of metropolitan areas was held down because they all shared in a greater migration from non-metropolitan areas and in the prevailing higher rates of natural increase. But now the importance of both these shared sources of growth is declining sharply, and without this cushion the migratory exchanges among metropolitan areas come to dominate the evolution of the national pattern of population distribution. With these common sources of growth much diminished, the relative variations in metropolitan growth rates will be greatly increased.

The magnitude of inter-metropolitan cross-flows is little realized. Although the 1970 Census data is not yet available, data for 1955-1960 can give a sense. In that period, metropolitan areas as a whole gained 1,220,000 through net domestic migration.[1] But this figure is the result of 12,400,000 departures from metropolitan areas and of 13,600,000 arrivals.[2]

---

[1]Migration from abroad to metropolitan areas was 1,550,000 in that period.

[2]These figures include movements among metropolitan areas and between these and the non-metropolitan remainder.

Everyone knows that the San Francisco area is a favored destination. Yet if departures from it had increased by only 15%, or arrivals decreased by 13%, San Francisco would have experienced no net migration. Metropolitan Los Angeles had a net migration of 442,000; but this was the net result of 1,653,000 arrivals and departures. Pittsburgh had a net outmigration of 67,000, but 113,000 moved into Pittsburgh. Philadelphia's 12,000 migratory gain was less than 2% of the cross flows of that area. Comparable figures apply to small metropolitan areas, from Saginaw, Michigan, to Savannah, Georgia.

Obviously, relatively small variations in the rates of arrivals or departutes can produce massive changes in the net figures. If the rate and direction of movements can be affected even slightly by public policy, there can be substantial control of the rates of population growth and decline of the various areas (expecially if natural increase continues to decline), and hence of the national pattern of population distribution.

In the last few years newly available data, computer capacity, and governmental policy interest and funding has led to a more active examination of these flows, to some preliminary answers, and to several questions and surprises. In the next pages I will tell of some of these findings and of their policy implications, focusing on two questions: where people go, and what affects the rate at which they leave. I will first present some of our statistical findings at Berkeley, and then supplement them with those of other recent investigations.[3]

Where people go.

People head toward a destination with a frequency not quite proportional to the size of it, just as the number who leave a metropolis is slightly less than proportional to its size. This means that there is no inherent migration bias toward larger metropolitan areas: about the same number of people will go from a large metropolitan area to a small one as will leave the small one for the large one, all other things being equal.

But all other things are not equal. Migration flows are attracted more than proportionally to the per capita income of an area, and since larger metropolitan areas tend to have higher average incomes, they tend to receive more than proportionate flows. Yet our findings clearly indicate that this is the effect of income differences, rather than of an inherent tendency toward larger areas. Small metropolitan areas with high average incomes will have positive migration balances with respect to large metropolitan areas of lower incomes.

---

[3]This work is still in progress, supported by the Economic Development Administration of the U.S. Department of Commerce. A somewhat fuller report of our preliminary findings is available in W. Alonso, "The System of Inter-metropolitan Population Flows," Working Paper 155, Institute of Urban and Regional Development, University of California, Berkeley, August 1971, and will appear in the report of the National Commission on Population Growth and the American Future.

We find too that distance reduces the attraction of a destination and that good climate enhances it. We do not find that unemployment affects the rate of arrivals, but there is a negative feedback such that, as the competition for the attractions and opportunities of a metropolis increases, many potential inmigrants appear to be discouraged and go elsewhere. This self-regulating reduction of inmigration probably accounts for the unimportance of unemployment.

These findings suggest some simple rules of thumb for a policy of growth centers in small or middle-sized metropolitan areas. First, it appears that income opportunities are far more important than size. In other words, high wages have more pull than the number of jobs. Second, that the location of the center should be near the population to be attracted. A metropolitan area of 200,000 which is 100 miles away has as much pull as one of 2,000,000 which is 1000 miles away. This applies not only to depressed and underurbanized areas, but also to attempts to control the growth of large metropoles. Growth centers which are to serve as countermagnets should be near the large metropoles where population pressure is to be relieved. We already see evidence of this process in the rapid growth of the smaller metropolitan areas within the national megapolitan constellations.

Third, that because of the negative feedback (and because outmigration increases with inmigration, as will be discussed below), attempts to have centers grow very rapidly will be grossly inefficient as the rate of effective attraction will increase far more slowly than the inducements offered. And fourth, that a national growth centers policy must be mindful of the interplay between size, rate of growth, level of inducements, and the number of centers. Too few centers brought along under hot-house conditions will be wasteful, but a policy of too many centers with lukewarm support would be ineffective.

The rate at which people leave.

My tests on the 1955-1960 migration data show that the rate of outmigration decreases slightly with metropolitan size, which is to be expected because of the greater range of opportunities larger places afford. Whereas a warm climate is attractive, climate appears to play no role in the rate of leaving. In other words, migrants head toward warm climates, but do not appear to flee from cold ones. The rate of outmigration correlates strongly with the proportion of young people. This is to be expected, since it is mostly the young who move. The rate also increases with previous growth. This confirms the well-documented finding that migrants into an area are far more likely to move again than natives. The data also shows that metropoles which have a large number of relatively near-by attractive destinations have a higher rate of departures, and we have called this a 'pull out' effect, which appears to have gone unnoticed in earlier studies.

None of these relations are suprising. The surprising finding is that neither low income nor high unemployment appear to correlate with outmigration. Indeed, higher incomes seem to increase outmigration. It is the absence of

these traditional "push" variables which is startling and has profound policy implications. It almost seems as if a metropolitan area emits emigrants like some radioactive body, proportionally to its size (adjusted for age and proportion of habitual movers), but essentially unaffected by such environmental conditions as its economic well-being.

The implication of such a finding, if true, would greatly change the basis of national policy toward distressed areas. A cornerstone assumption of that policy, in this country as in others, is that economic distress leads to heavy outmigration, and that this inflicts upon the area of origin grave social costs as well as on the area of destination. Policies of investment in depressed areas have been based in large measure on an attempt to expand local economic opportunities and prevent economically-forced outmigration. In effect, they aim to provide people the freedom to stay in their own region.

But if my findings (and similar ones wuch as those of Lowry[4] and Lansing[5]) are correct, the picture changes radically. What seems to be heavy outmigration is then only a figment of statistical aggregation. The negative net migration of a declining area is the result of a shortage of inmigrants, not of an excess of outmigrants. But no human being is a net migrant: people can come or they can go; the net is the arithmetic difference. Thus programs of development would not have the effect of retaining the original inhabitants, but rather that of increasing the inflow of outsiders in to areas which typically already suffer from a labor surplus. Since population retention has played such a central role in national policies toward distressed areas, if this finding is correct we must rethink the objectives and the instruments of most regional policy.

Yet, precisely because of its potential policy importance, this finding should not be simply taken at face value. A number of recent authors have quite properly addressed themselves to determining whether Lowry's findings and mine are a statistical mirage. Unfortunately, this work is at present dispersed, largely unpublished, and not systematically related to policy issues. Although I shall review these studies critically, it is not to downplay their importance or their quality, but rather to convey some flavor of the difficulty of the inquiry and the importance of its implications for policy.

Lee D. Olvey, in a 1970 Harvard Ph.D. thesis[6], objected that Lowry's failure to find a push from unemployment was the result of his examining this relation after the outmigration had taken place; Olvey argued that the real cause for outmigration would be the unemployment that would have occurred

---

[4]Ira S. Lowry, Migration and Metropolitan Growth: Two Analytical Models, (San Francisco: Chandler Publishing Co., 1966).

[5]John B. Lansing and E. Mueller, The Geographic Mobility of Labor, Ann Arbor, Survey Research Center, Institute for Social Research, 1967.

[6]Lee D. Olvey, "Regional Growth and Interregional Migration: Their Pattern of Interaction" Ph.D. Thesis, Department of Economics, Harvard University, 1970.

in the absence of outmigration. He constructed a measure of "prospective unemployment" for metropolitan areas, and found that it correlated very strongly with outmigration. The measure, in effect, set the natural increase of the area, plus its inmigration, against its growth in jobs; that is, "prospective unemployment" was the gap between the increase in jobs and in the potential increase in the supply of labor in the absence of outmigration.

In spite of its common sense, the test is not conclusive, however, It may be that Olvey is simply relating outmigration to outmigration. Both in prosperous and distressed areas there is a close association between jobs and population. Thus, when the increase in jobs is taken out of what the population would have been without outmigration one is left with something very much like outmigration, almost by definition. It is thus not surprising that the prospective unemployment would correlate with outmigration, since the relation is largely tautological. On this basis, it is very doubtful that this evidence shows that prospective unemployment causes outmigration.

Olvey also found that low wages accelerated outmigration; however this relation held only in one of his statistical experiments, and not in the others. Thus this relation is statistically weak.

Charles E. Trott's work[7] in the Bureau of Economic Analysis of the U.S. Department of Commerce is of particular interest because he uses the Social Security Continuous Work History Sample and Office of Business Economics Areas rather than Census materials and metropolitan areas, and because he looks at variations with age and with race. His most interesting finding is that outmigration appears to relate not to wages or income as such, but to the ratio of the area's expected wages (if national wage levels applied to the area's industrial structure) to the area's actual wages. That is to say, that the propensity to migrate depends not on low actual wages, but on how these wages compared with the national averages for that industry. He also found a relation of outmigration to the ratio of new labor force (new entrants into the labor force plus immigrants) to new jobs in the area, a measure similar to Olvey's "prospective unemployment". Thus, wages and unemployment as push factors re-enter the picture. However, in my opinion, the finding is not a very certain one yet. It is beyond the scope of this paper to enter into detailed statistical arguments, but let me point out that these variables of expected wages and expected unemployment are elaborately synthetic ones, and that of necessity they require many rough approximations and implicit assumptions.

Trott's other findings are also of interest. He finds, for instance, that the proportion of workers in manufacturing accelerates outmigration.

---

[7]Charles E. Trott, "Differential Responses in the Decision to Migrate," Bureau of Economic Analysis, U.S. Department of Commerce, mimeo, November, 1971.

This creates some question in the consideration of national programs of job creation in distressed areas, which have stressed manufacturing development. In another paper[8], Trott finds a surprising negative association of outmigration with the ratio of actual to expected employment. This seems to run counter to the prospective unemployment findings. He also finds that outmigration in fact increases with an earlier high rate of employment increase, a finding that may have to do with the foot-looseness of earlier migrants.

Edward Miller, of the Office of System Analysis and Information of the U.S. Department of Transportation, in two very recent and as yet preliminary studies focuses in some additional variables, principally the accelerating effects of education on outmigration and the roles as migrants of those born within or outside a state. (The outmigration rate of the latter is nearly three times that of those born in the state.) He finds[9], for instance, that (using Census data) there is no direct association of income and outmigration. But when account is taken of some other variables, principally the number of out-of-staters and the level of education, low income becomes strongly associated with outmigration. Similarly, outmigration and job-formation at first appear unrelated, but job-creation appears to reduce outmigration significantly if other factors (such as inmigration and education) are taken into account. The difficulty is that in the statistical forms in which job creation becomes significant, income stops playing a role. Yet when BLS data is substituted for Census data for some variables, both employment growth and high wages reduce outmigration as expected. But in none of his tests does unemployment register significantly.

Miller's other study[10] is of particular interest because he focuses on the outmigration of those born within a state. They are presumably the clients for those policies which aim to make it possible for people to earn a living without leaving their communities. Using Census data, he finds almost the perfect relation supporting the push hypothesis of migration. Outmigration is slowed by growth in jobs, is accelerated by education, reduced by income (although this variable is weak), and increased by unemployment. Using BLS data he finds about the same relation, except that low income is greatly strengthened as a push for outmigration; but unemployment drops out as a factor.

Let me try to put simply my conclusions from this brief review of where we stand on the matter of what determines outmigration. There is ample evi-

---

[8]Charles E. Trott, "An Analysis of Outmigration," O.B.E., U.S. Department of Commerce, mimeo, August 1971.

[9]Edward Miller, "Determinants of Outmigration - Why Study Outmigration" Office of Systems Analysis and Information, U.S. Department of Transportation, November, 1971.

[10]Edward Miller, "Out-migration Rates for Those Born Within a State," Office of Systems Analysis and Information, U.S. Department of Transportation, November, 1971.

dence that there is a strong mechanism at work here, but because the various factors are interrelated we cannot set out cause from effect. Until we understand these processes better, there is a good chance that our policies and programs may not only be wasteful or ineffective, but indeed counter-productive. Is there a push from low income or unemployment? The evidence here is murky, some finding one, some the other, some rather abstract versions of one or the other, some neither, and even in some cases a reverse effect. If there is no push from low wages and unemployment, present policies are wasteful in terms of retaining the present population, and possibly counterproductive if, by encouraging inmigration, they accelerate outmigration.

But if there is a push, is it a push of too few jobs or of low wages, or both? Most of our present programs appear designed to develop jobs rather than to raise wages, and thus implicitly to vote for a job-gap theory. But the evidence is at least as strong that it is an income-gap which is at work. Yet trying to raise wages in an economically distressed area is a chancy proposition, because perhaps their most usual principal comparative advantage is low-priced labor. If the income-gap is one in terms of money income, an area might raise its average wages by attracting high wage industries at wages slightly lower than the national average for that industry. But if the income gap is in terms of relative wages (as in Trott's findings), this would be just the wrong policy. In the long run, policies aimed at raising incomes must concentrate part of its efforts in human resource development (a vague term which means at least improvement of educational levels) to produce a people who is worth better wages and perhaps even able to generate some of its own develoment. Although education as a policy would appear to be like motherhood, it is a two-edged policy; for every study which has looked into it agrees that higher education leads to higher outmigration.

This brief exploration of the relation of migration flows and regional policy has obviously dealt with only a few of the many aspects of the realities of regional development and of the policies addressed to these issues. Yet the movement of people is what shapes the distribution of the national population and thus must be a central theme of any national territorial policy. I have tried to show that we are already learning some things which are directly useful for policy formulation, and that we are at the same time discovering those questions which must yet be answered to have effective and well-aimed policies. It is clear that we must continue the present national intensification of research on these matters, and that we must try to relate better the findings of such studies to issues of policy and of program. It is not only that we do not yet know what works, but that we do not yet know clearly what we want.

Finally, I want to raise an emerging but largely unrecognized situation in which these issues arise. This is the prospect that a great many metropolitan areas will have stable or declining populations a decade or so from now. Not just their central cities, but these areas as a whole. This is inevitable if we begin to approach zero population growth nationally. The

lively interchange of migrants among metropoles is certain to continue. If the total population is fixed, gains and losses must balance out. Those losing in the net through migration will also face lower birth rates because there is a demographic multiplier involved. Migrants are predominantly at the most fertile ages, and thus, if the national natural increase is near zero, areas that experience a net loss in the exchange of migrants will have the negative rates of natural increase since they will have a low proportion of the young.

This will be a new situation. Previous high rates of urbanization and of natural increase has insured that with rare exceptions all metropolitan areas grow in population. The few that have lost are economically troubled ones. But, unless there is a new reversal in birth rates, we will have dozens of reasonably prosperous metropolitan areas with shrinking populations. There is at present no serious work being done to identify the problems or opportunities that such a situation will present, and thus we are again likely to be caught unprepared. Indeed, virtually all of existing economic theory and empirical calibration of relations are geared to the phenomenon of growth. Yet the path of decline in population is not likely to be sliding back along the same track as that of previous growth. It is likely to follow a different path.

If economic growth continues at the same time that metropolitan population is shrinking, while still involving a substantial exchange of migrants, several interesting possibilities are open. The proportion of children in the population will decline more sharply than for the nation as a whole, and thus the fiscal pressure of educational costs at present standards will abate. This will be reinforced by the higher labor participation rate of the population because of the more top-heavy age structure. Welfare costs are not likely to be very different because even today migrants are no likelier to go on welfare than natives. New investment in urban infrastructure is likely to decline, since there is no need to accommodate future population. Similarly, the demand for new housing will be reduced. In general, the construction industry should decline.

But on the other hand, the relative scarcity of labor suggests that there may be a tendency to substitute capital for labor, while the lightened load of education and added infrastructure costs may increase the propensity to save and thus increase the supply of capital available. The maintenance and use of the existing capital stock may acquire new importance. The housing stock in particular, is likely to see at the same time a deflation of values in central city property as suburban land. This may create sharp problems of adjustment. At the same time, the filtering process may deliver proportionately more housing to the lower income groups in the central cities so that they will be better off in that respect. Yet this may present problems in terms of their access to jobs from their central locations and of the social integration of central city and suburbs because of the greater relative costs of new construction. Similarly, the rate of abandonment may increase.

More subtle questions are also likely to be important. For instance, what will be the consequences of such a situation for the economic rise of minorities? Recent studies show that their economic status is considerably higher in larger places[11], and a recent opinion survey[12] indicates that they are more in favor of contined growth than the white population. Similarly, what will this do to the careers of the areas' young people? On the one hand, there will be fewer opportunities and more entrenched older persons, on the other hand there will be far fewer young people. And will social tranquility be greater? The crime rate appears lower in today's declining and low growth metropolitan areas[13], but this may be the result of the low proportion of the young, who share disproportionately in crime. Most generally, will there be in these areas increased socio-economic integration as a result of reduced pressures, or will there be more rigidity in social stratification, less innovation, and more cut-throat competition?

The point here is not that the issues are likely to be dramatic. Indeed, prosperous population stability or slight decline seems at this distance like a comfortable middle age. Yet even that can have its problems, although they may not be tragic. Given the likelihood that this will be the future of dozens of our metropolitan areas and tens of millions of our citizens it seems important that we do some anticipatory research and planning. Our present theories and rhetoric deal primarily with size. This derives on the theoretical side from the static equilibrium bias of most social science, and on the rhetoric side with the traditional utopianism of ideal size. But size and growth must be distinguished. Such theory and traditional rhetoric as we have on change deals with growth, since this has been our experience for two centuries of national life. A new rhetoric has arisen against growth, but it deals primarily with the need for stopping growth. The likelihood appears to be that in many cases population growth will stop or reverse spontaneously, and we should now anticipate the problems and opportunities of this in the near future.

---

[11] For a useful review of this and related matters, see Peter A. Morrison, "The Impact and Significance of Rural-Urban Migration in the United States," The Rand Corporation, Santa Monica, P-4752, March, 1972.

[12] Unfortunately, this recent national survey at this writing still is not to be cited directly.

[13] Edgar Rust, "Metropolitan Non-Growth," Department of City and Regional Planning, University of California, Berkeley, March 1972.

SUMMARY

Concluding Remarks of William W. Blunt, Jr.

William Alonso has asked the question, "Is regional development worth attempting?" His remarks are well worth considering. We certainly should ask ourselves whether it is desirable to have a Federal policy to influence regional growth, and to influence the distribution of the population among the regions. This is a key question that brings us directly to the point: What is it that we are trying to accomplish, and what are the effects of our efforts? Consideration of these questions provides us with a framework for increasing our awareness of the factors we are dealing with, and for guiding us in increasing the effectiveness of our program.

This conference has provided a valuable interchange among the research people, the EDA management and operating people, and other government offices. It has revealed, on the one side, the difficulties faced by the researchers in carrying out their work and in translating the results for those of us who must form policy and carry out the operating programs, and on the other side, the difficulties faced by those of us who are in the position of trying to understand the research results and put them to use. It has also given us an unusual opportunity to get the views of people from other agencies who are providing us with services or who are operating complementary programs. All of our efforts should be enhanced by the light we have been able to shed on our mutual problems and interests.

I believe we would all agree that the Conference has been very worthwhile and has given us a much needed opportunity to exchange points of view. I believe we should make it an annual event.